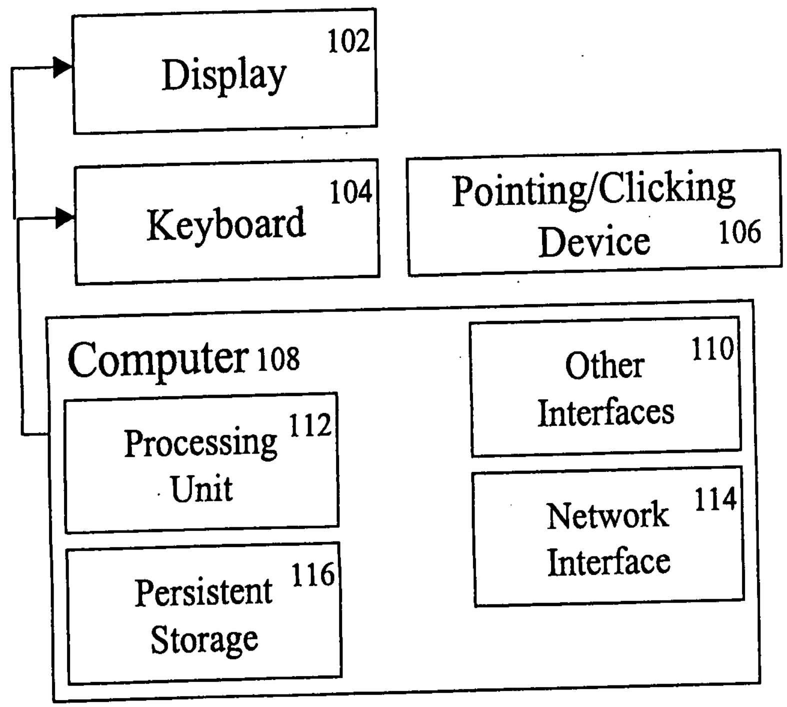
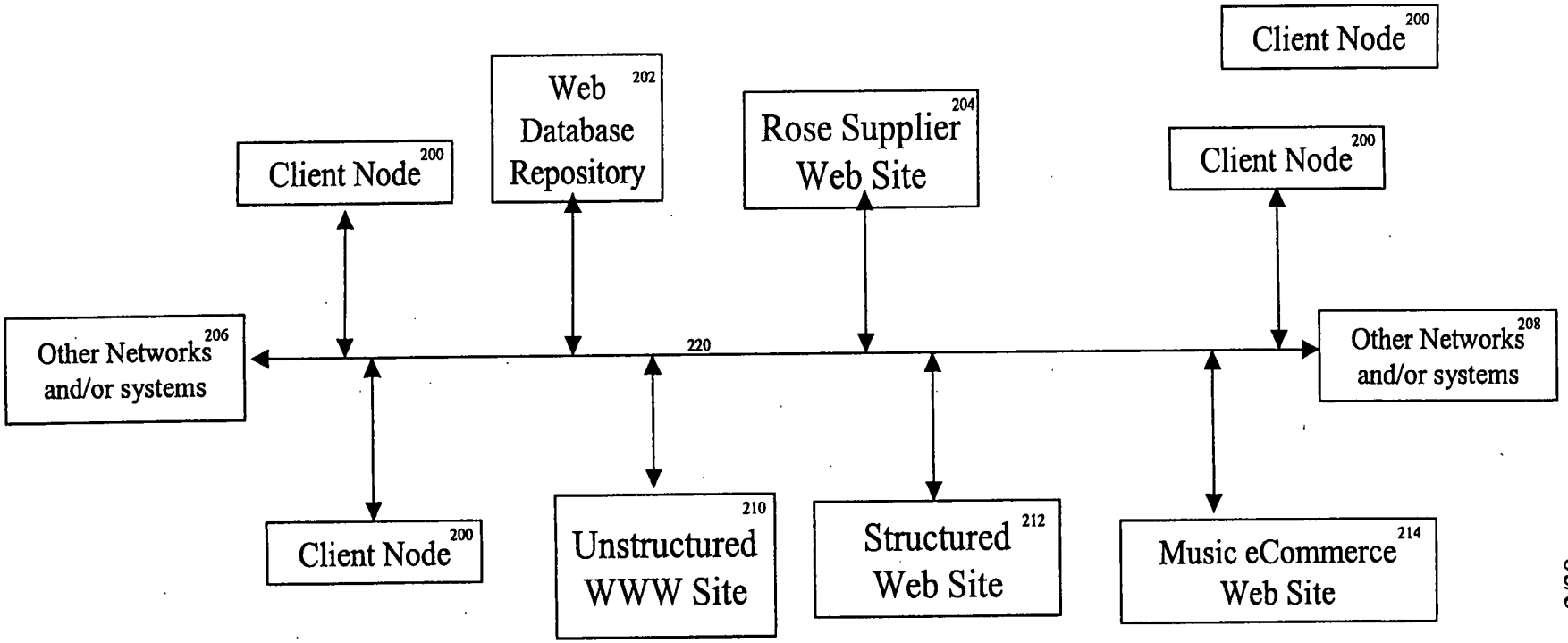
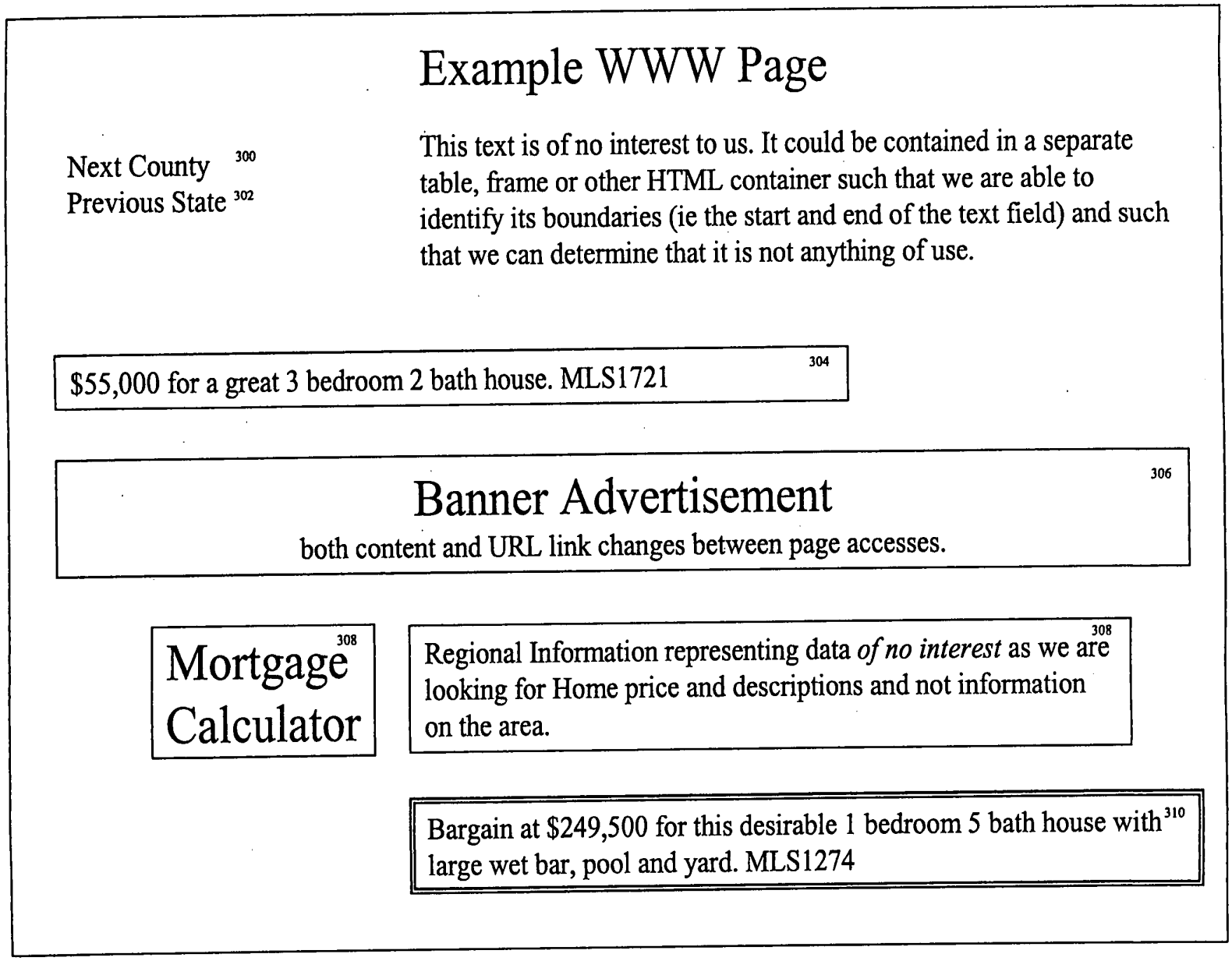


Computer Environment
Figure 1



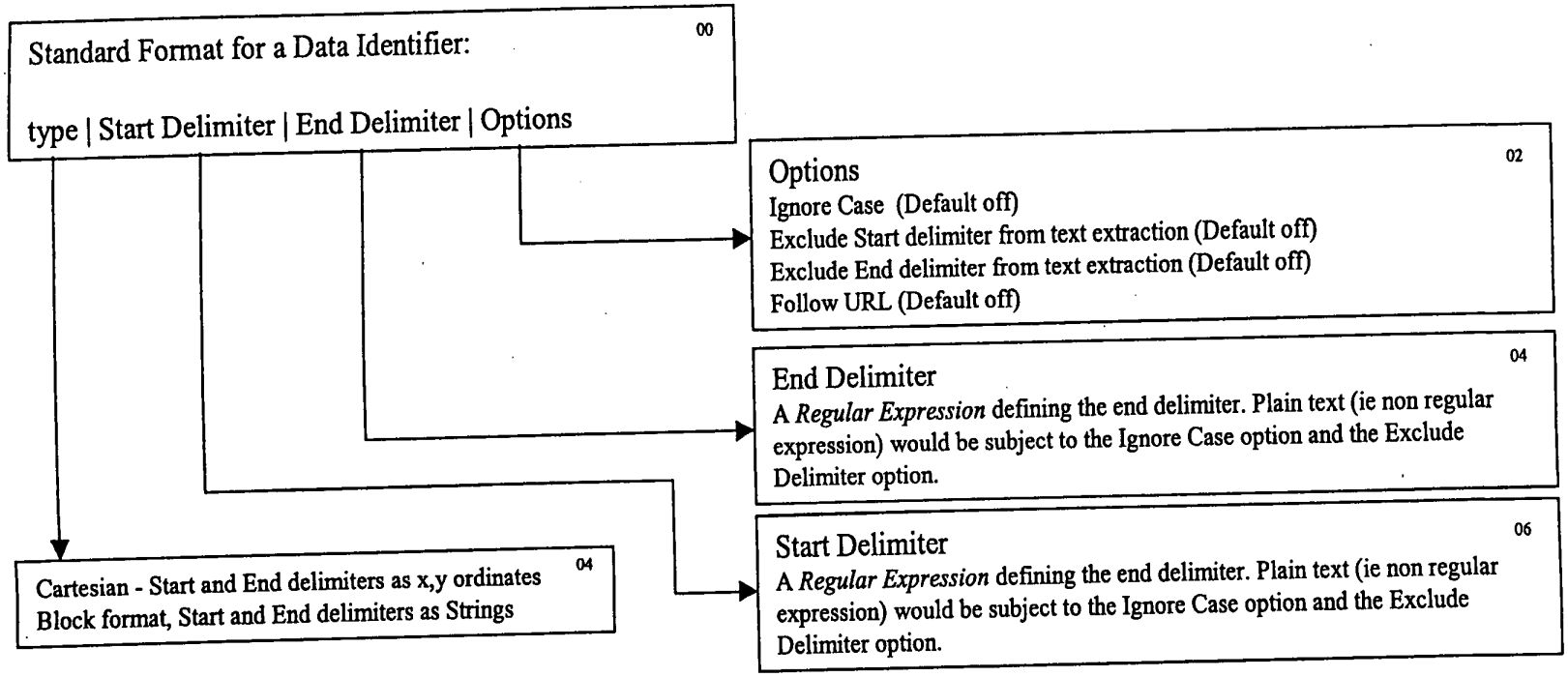
Shows a typical example embodiment where the invention is installed into a plurality of Client Nodes extracting data from a plurality of data repositories.

Figure 2



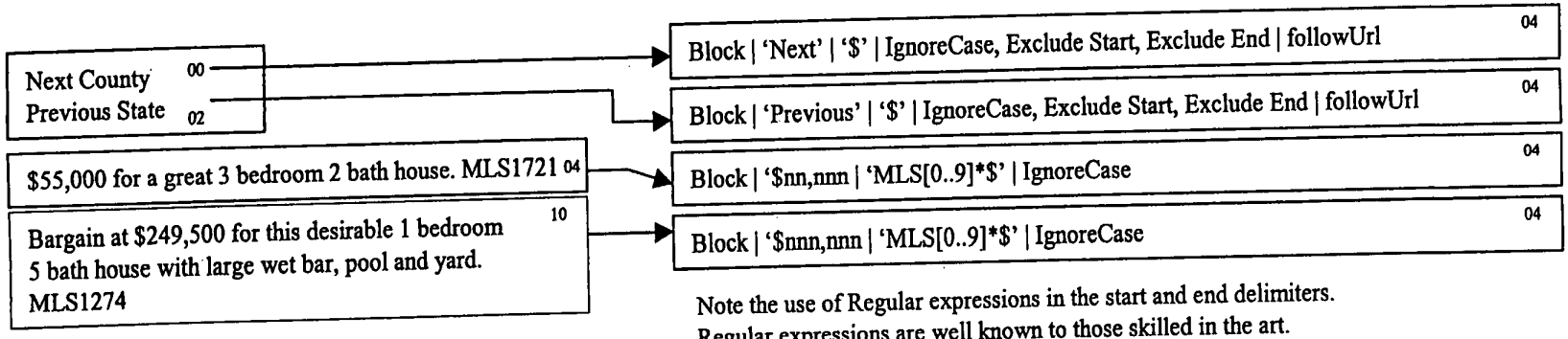
Example WWW page with data *of interest* and data *of no interest.*

Example WWW page
Figure 3



Text to Decode

Example Data Identifier

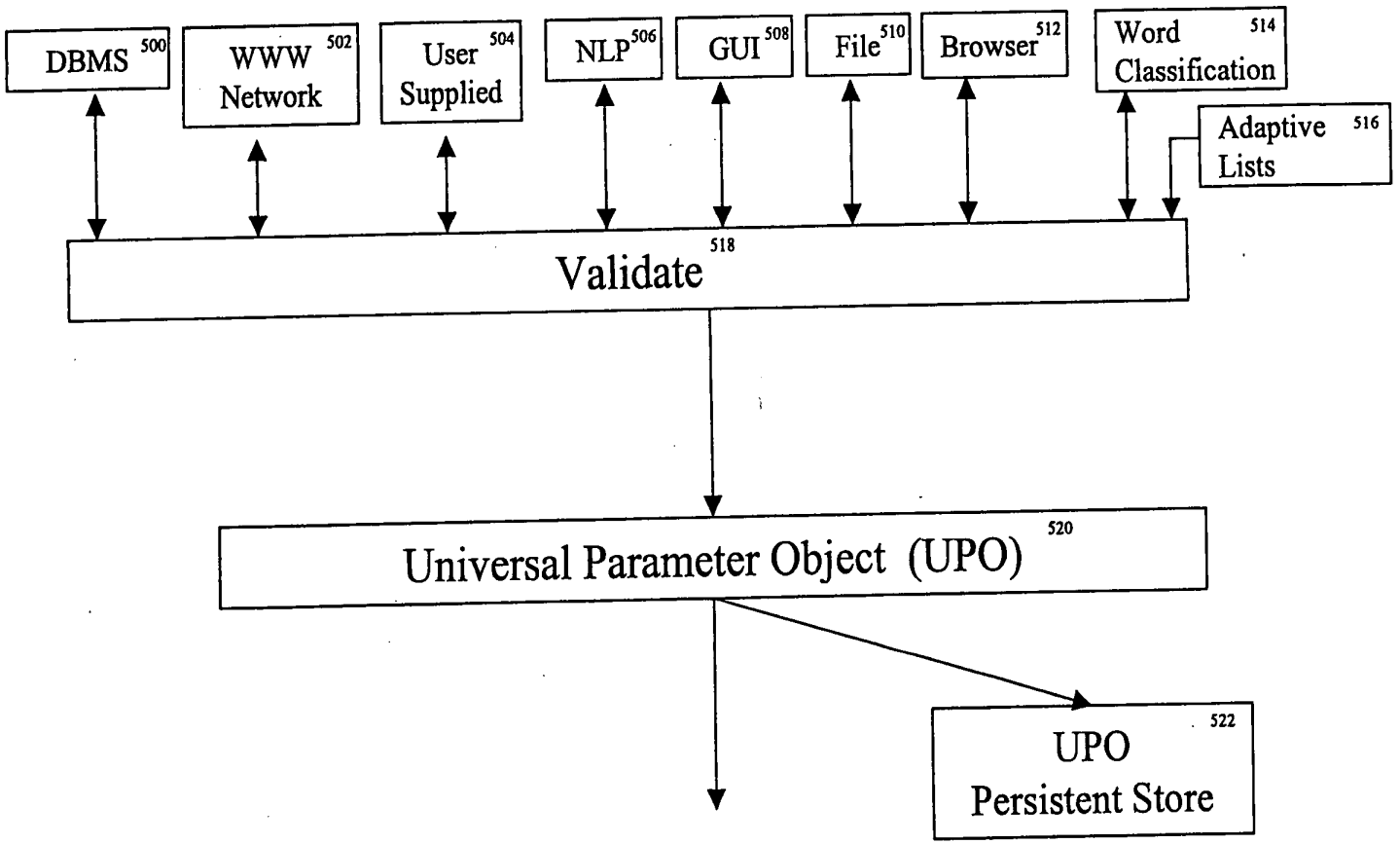


Note the use of Regular expressions in the start and end delimiters.
Regular expressions are well known to those skilled in the art.

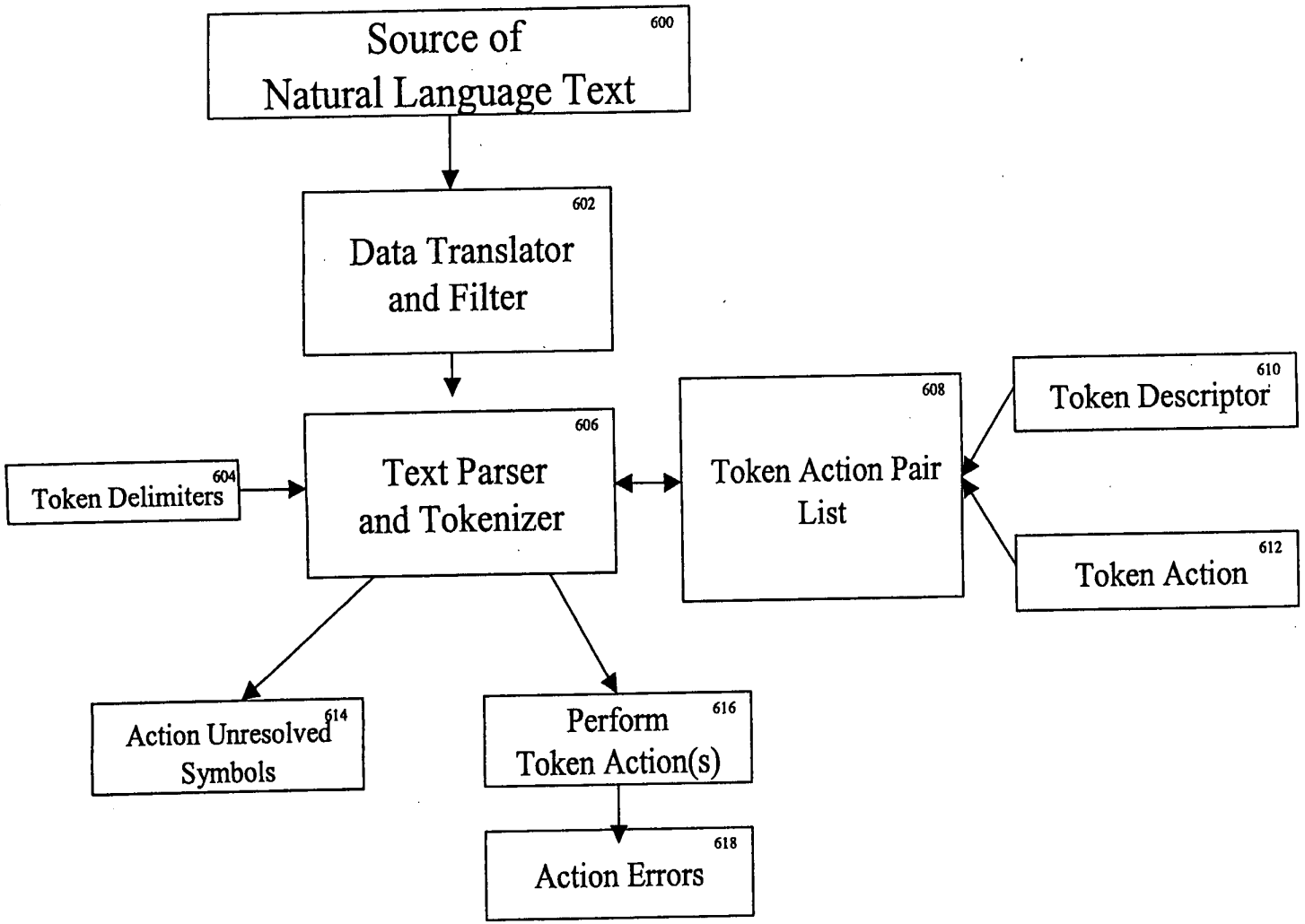
Example Data Identifier Usage

Figure 4

Universal Parameter Object (UPO)



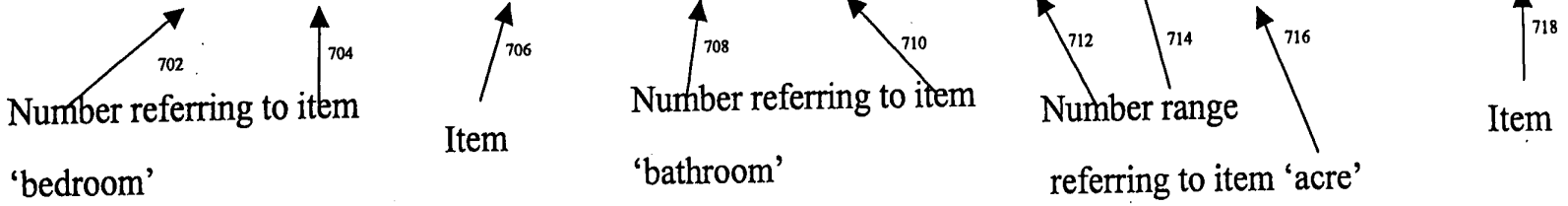
Universal Parameter Object (UPO)
Figure 5



Natural Language Processor
Figure 6

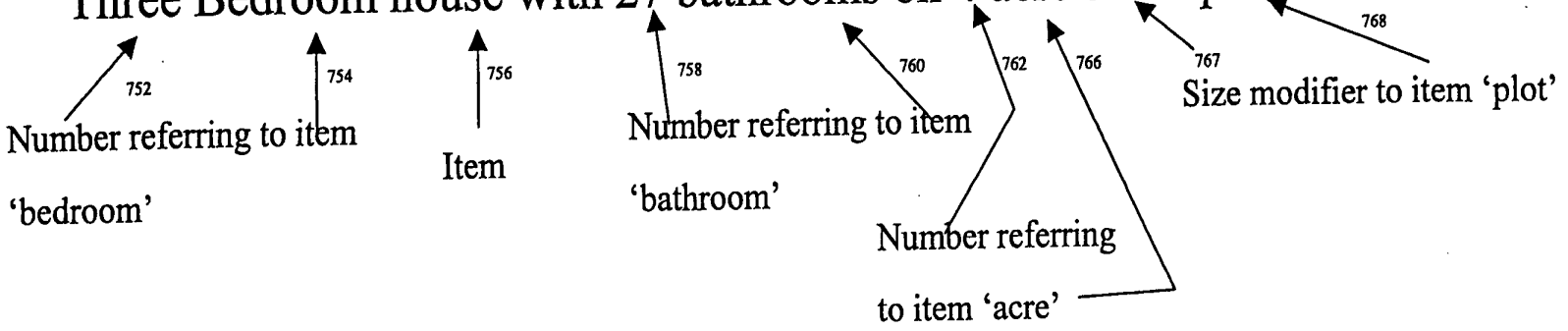
An example of Natural Language with parameters:-

700 Three Bedroom house with 2 bathrooms on 3 to four acre plot with view



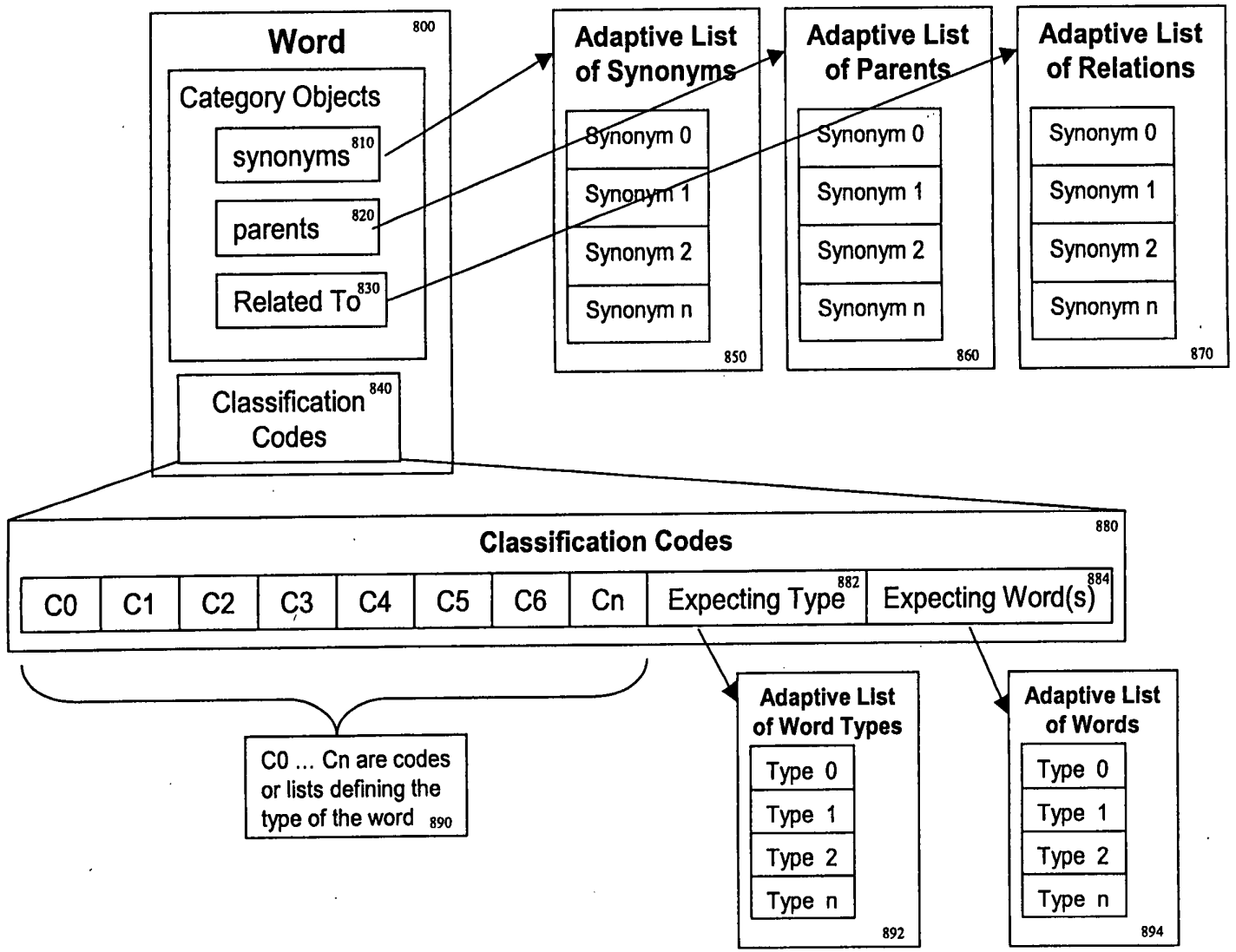
750 An example of Natural Language with improbable and conflicting parameters:-

Three Bedroom house with 27 bathrooms on 4 acre small plot



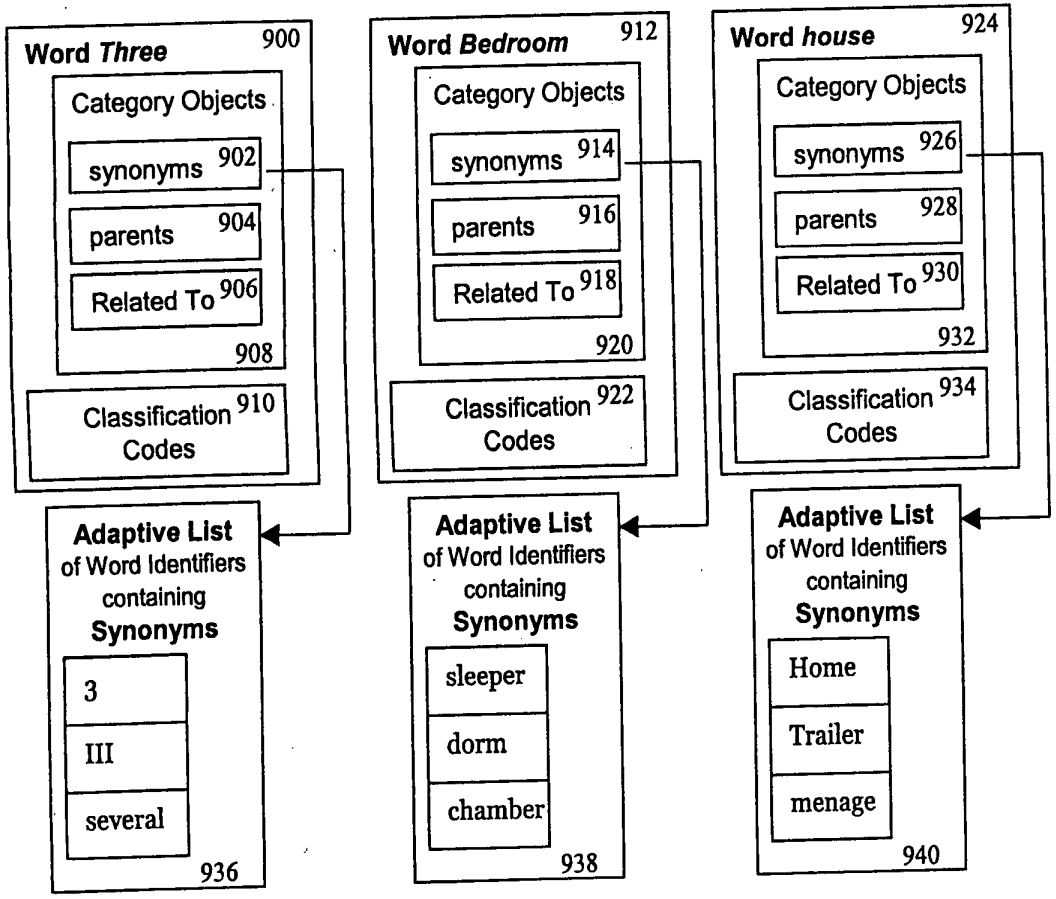
Examples of Natural Language
Figure 7

Word Identifier

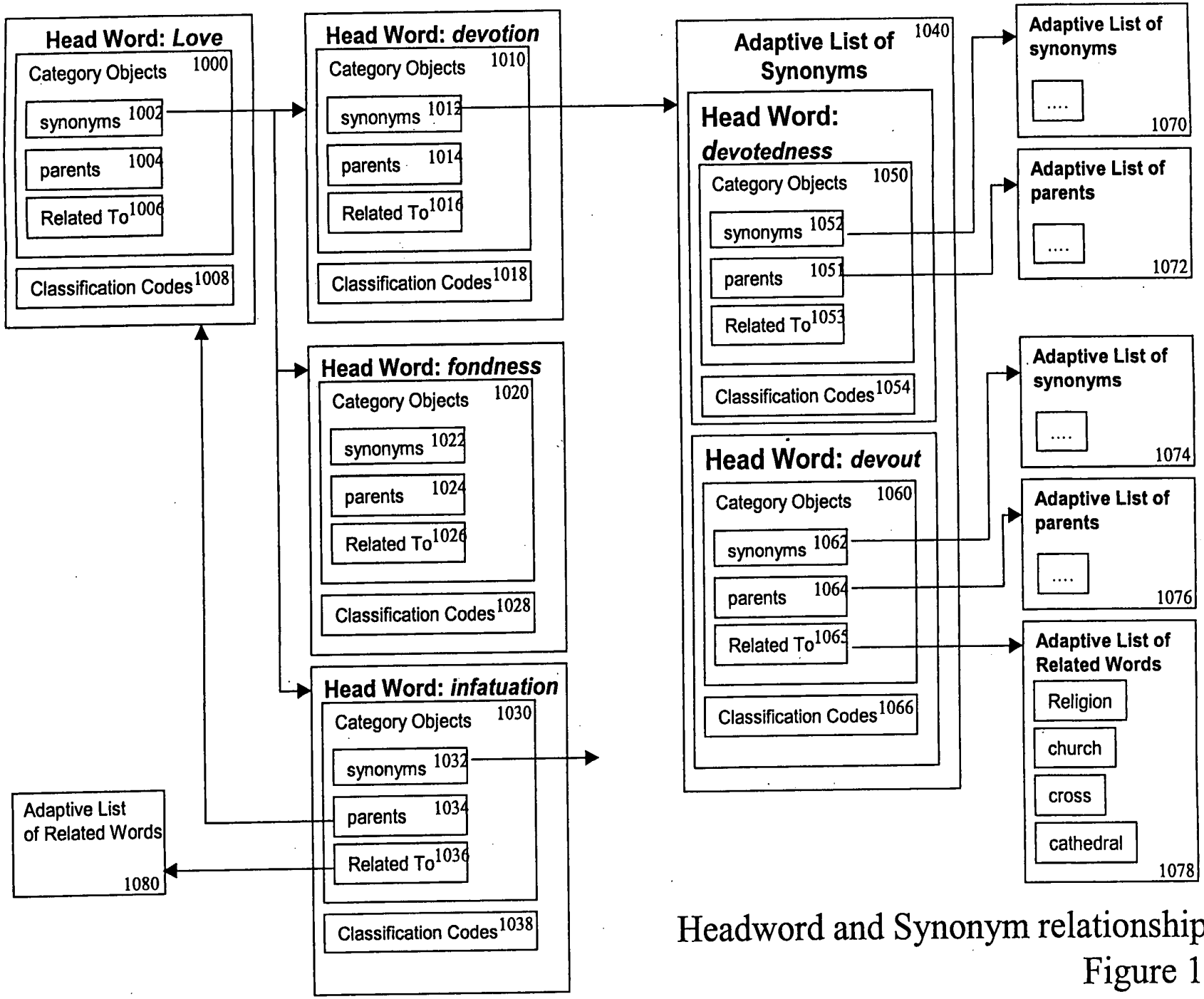


Word Identifier
Figure 8

Three Bedroom house with 2 bathrooms on 3 to four acre plot with view



Natural Language to Word Ident Mapping
Figure 9



Headword and Synonym relationships

Figure 10

Word Classification

Category Codes 1100									
C0	C1	C2	C3	C4	C5	C6	Cn	Expecting Type	Expecting Word(s)

Example word "cat" showing example categories it fits into. Such categories will vary between embodiments.

Target Word: "Cat"

Natural Category 1002 1110	Gender 1 1112	Species 1223 1114	Food Group 1000 1116	Function 5002 1118	Specific 0 1120	Specific n 1122

Example Natural Category Codes	Example Gender Codes	Example Species Codes	Example Food Group Codes	Example Function Codes
Flaura (2000)	None 0	Felis catus (1223)	Carnivore (1000)	Sleeps (5002)
Fauna (8000)	Female 1		Herbivore (2002)	Worker (6000)
Mammal (1002)	Male 2	Dogimus Woofus (2002)		
Insect (1003)				
Reptile (1004)				

The number, meaning and definition of categories will be dependant on the specific embodiment.

Word Classification Figure 11

Word Comparison

Category Codes 120

C0	C1	C2	C3	C4	C5	C6	Cn	Expecting Type	Expecting Word(s)

	Natural Category	Gender	Species	Food Group	Function	
Item 1: "Cat"	1002	0	1223	1000	5002	1212
Item 1 : "Tulip"	8000	0	7014	3007	9668	1214
Difference (distant match)	7000	0	5791	2007	4666	1216

Comparison of the codes for the words "cat" and "tulip". The difference is large as obviously a "cat" is an animal and a "tulip" is a flower.

1210

	Natural Category	Gender	Species	Food Group	Function	
Item 1: "Cat"	1002	0	1223	1000	5002	1220
Item 1 : "Lion"	1002	0	1227	1000	5003	1222
Difference (Almost exact match)	0	0	4	0	1	1224

Comparison of the codes for the words "cat" and "lion". The difference is very small as obviously a "cat" is closely related to a "lion"

1226

1230 **Difference** $\Delta_{c0} = {}^{0}I_{c0} - {}^{1}I_{c0}$

1232 **Difference** $\Delta_{c1} = {}^{0}I_{c1} - {}^{1}I_{c1}$

1234 **Difference** $\Delta_{cn} = {}^{0}I_{cn} - {}^{1}I_{cn}$

1236

$$Proximity_{T} = I_{c0} - I_{c1}$$

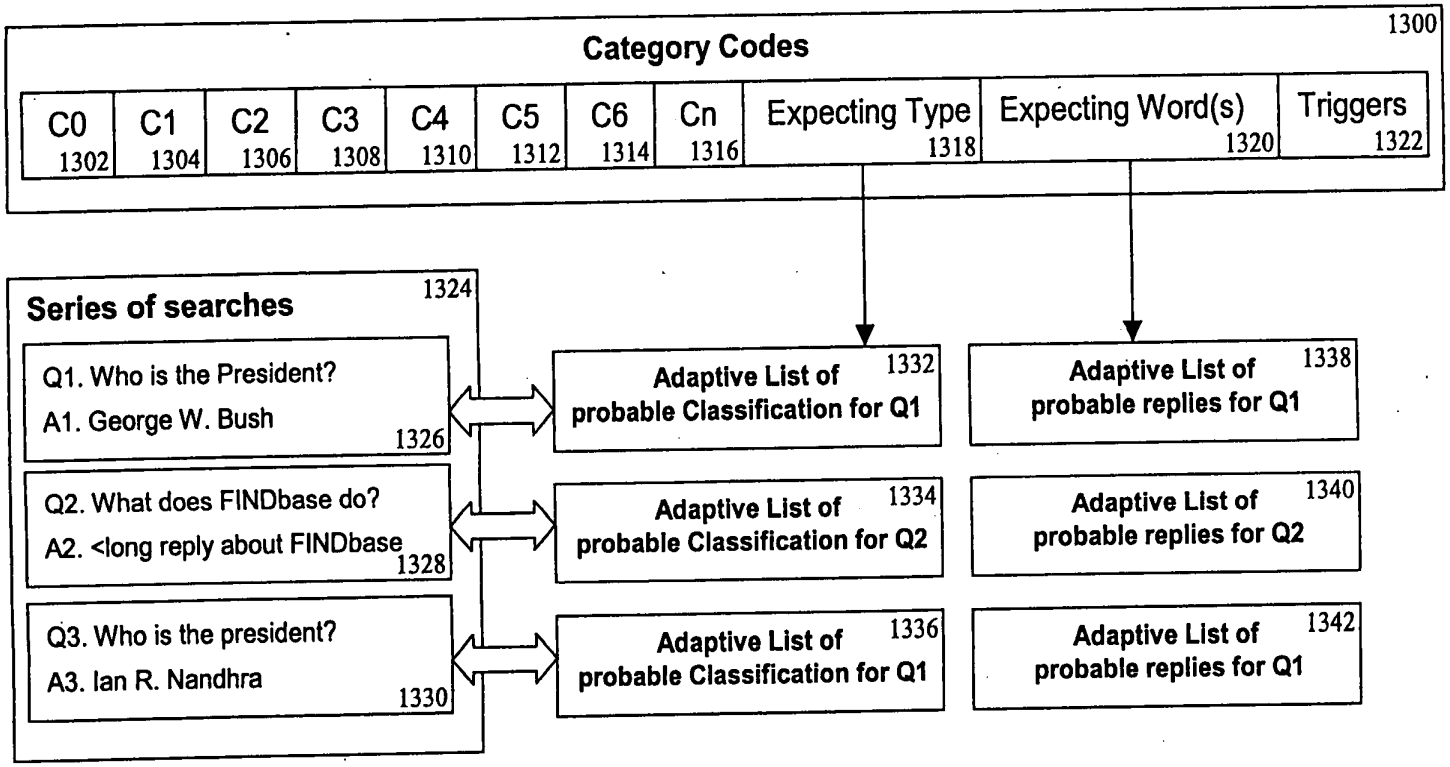
1238

$$Proximity_{cn} = \sum_{\Delta_{cn}}^{cx} \Delta_{cx}$$

Where cx and cn encompass a set or sequence of cells such as c0, c3, c4, c5, c9

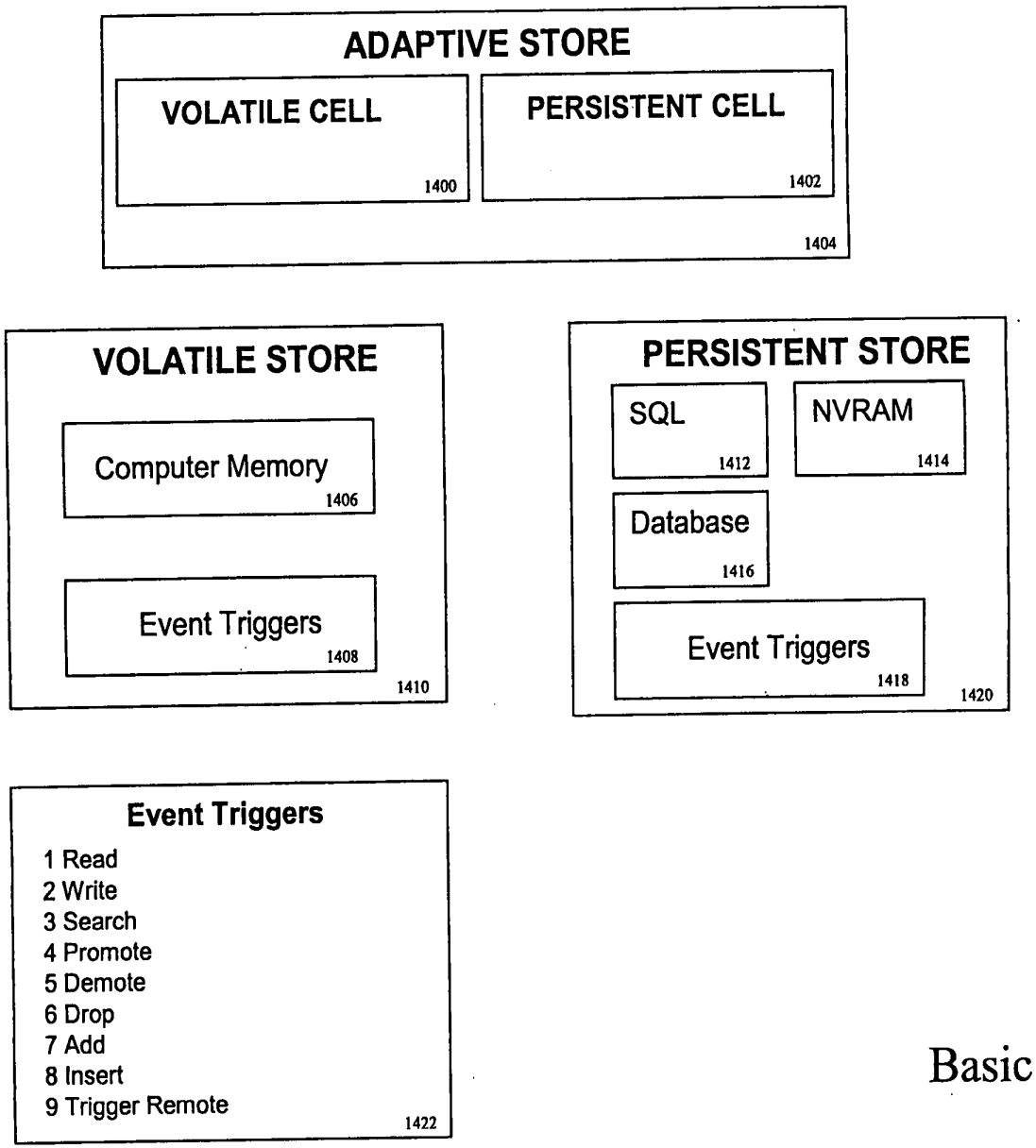
Word Comparison Figure 12

Word Expectation(s)



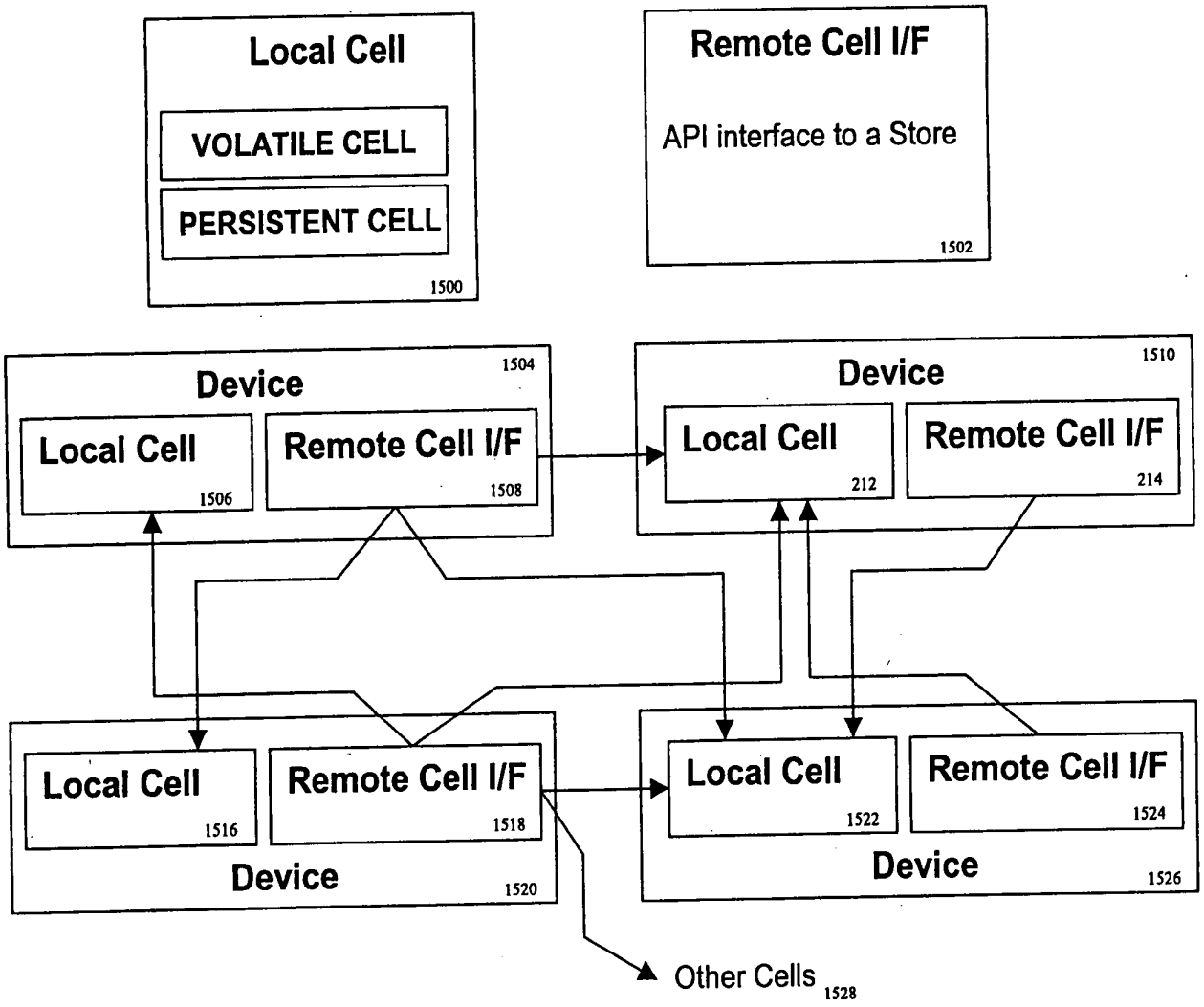
Word Expectations Figure 13

Basic Stores



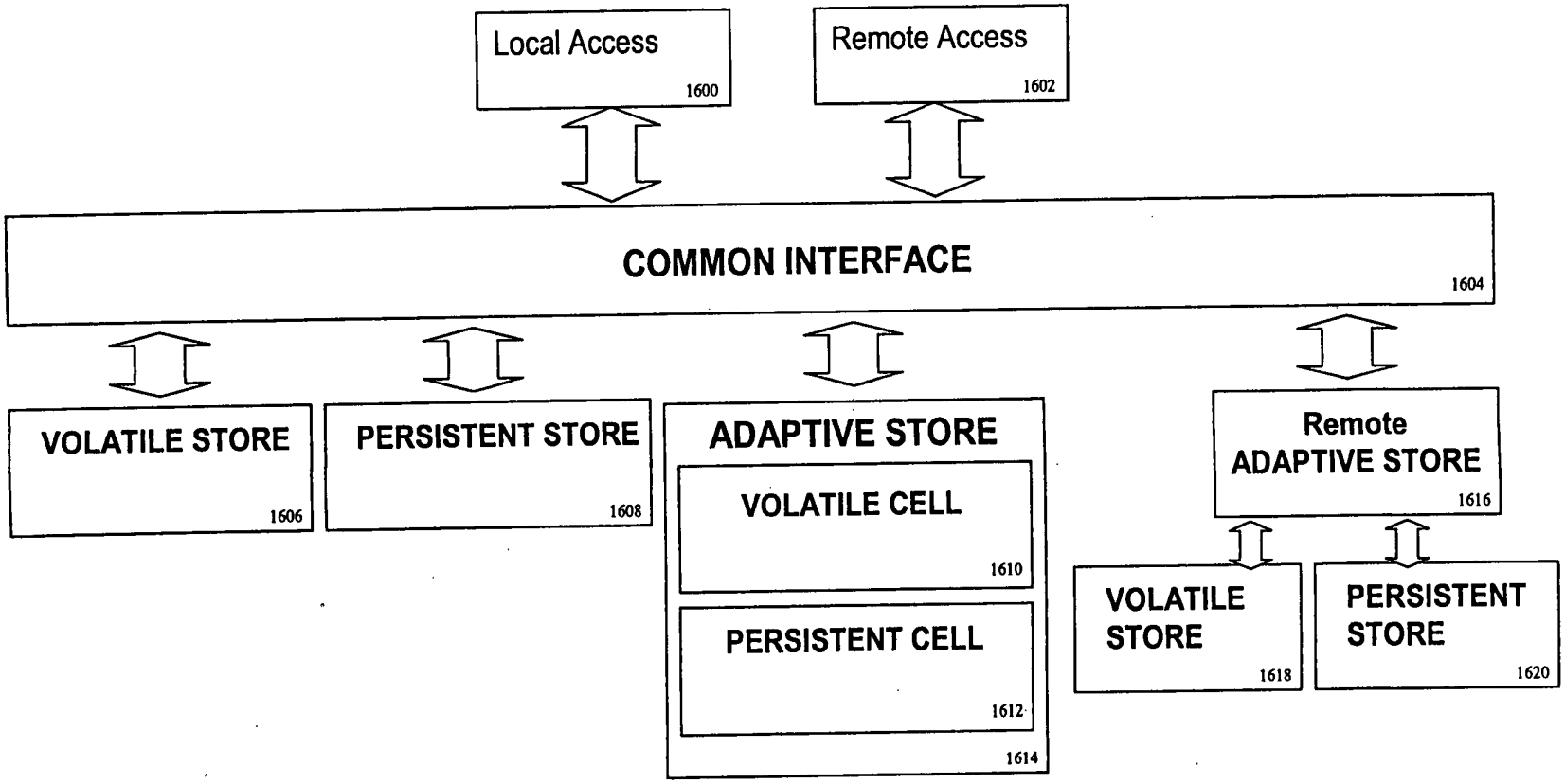
Basic Store Types
Figure 14

Storage Cells

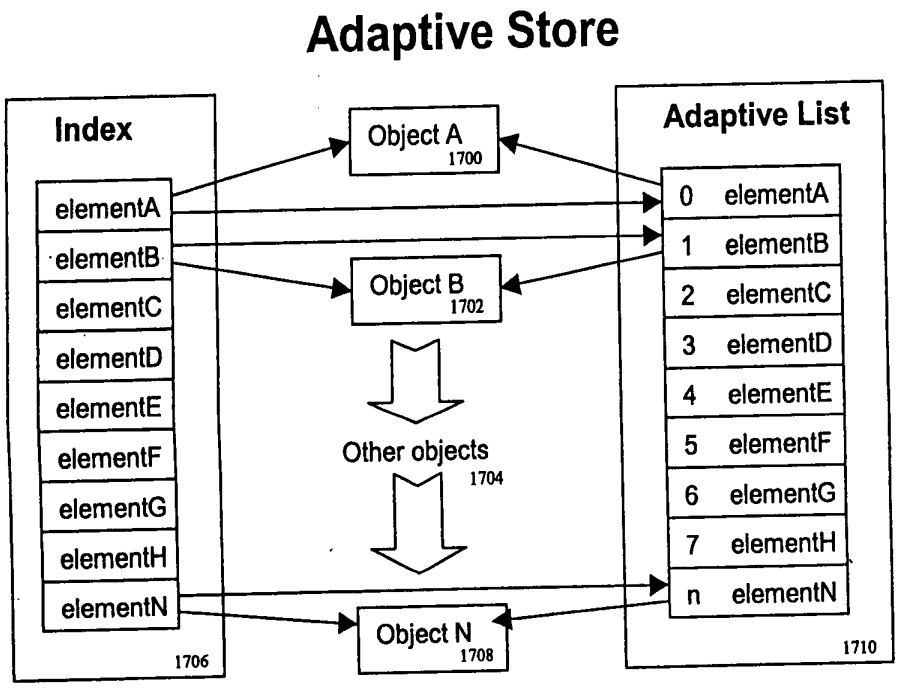


Storage Cells
Figure 15

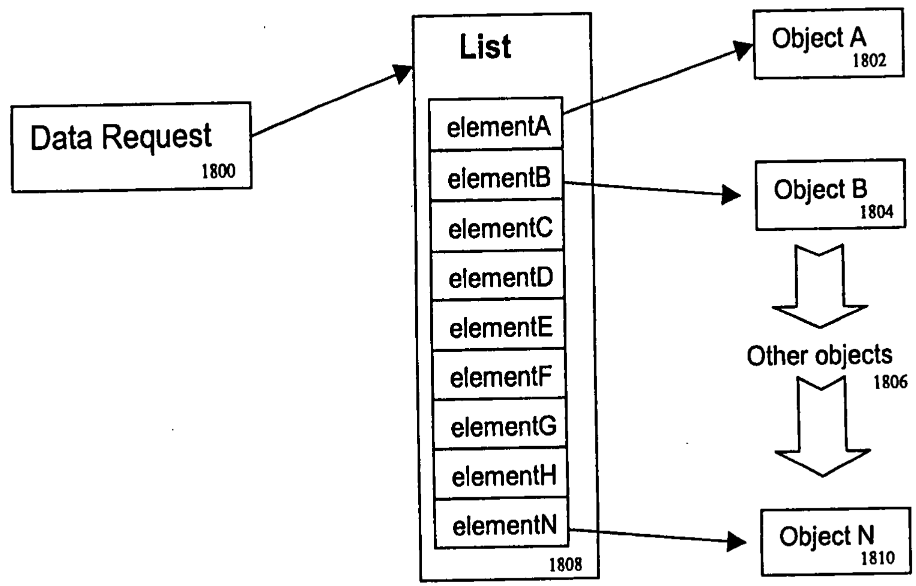
Storage API



Storage API
Figure 16

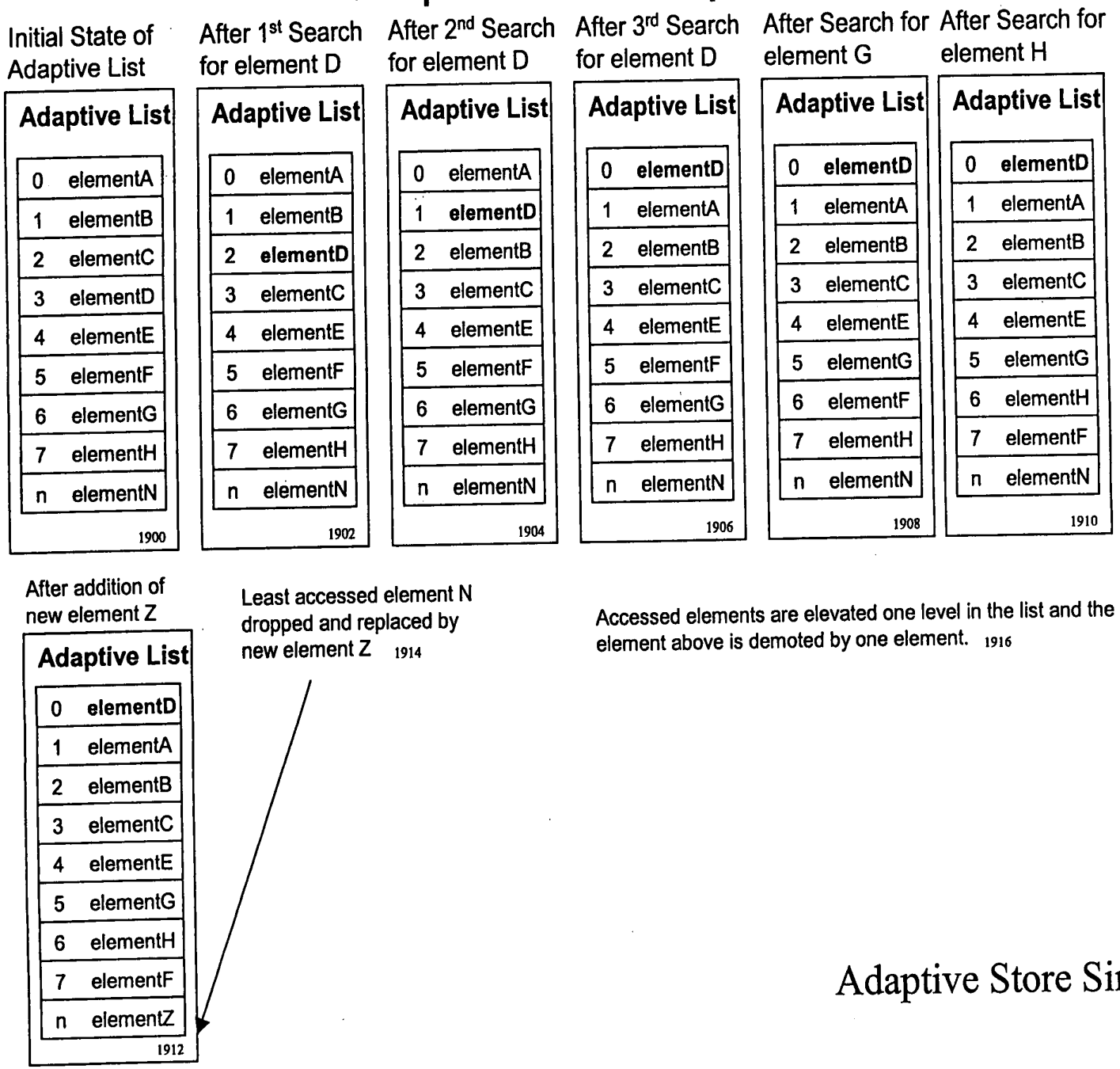


Adaptive Store Indexing and List relationship
Figure 17



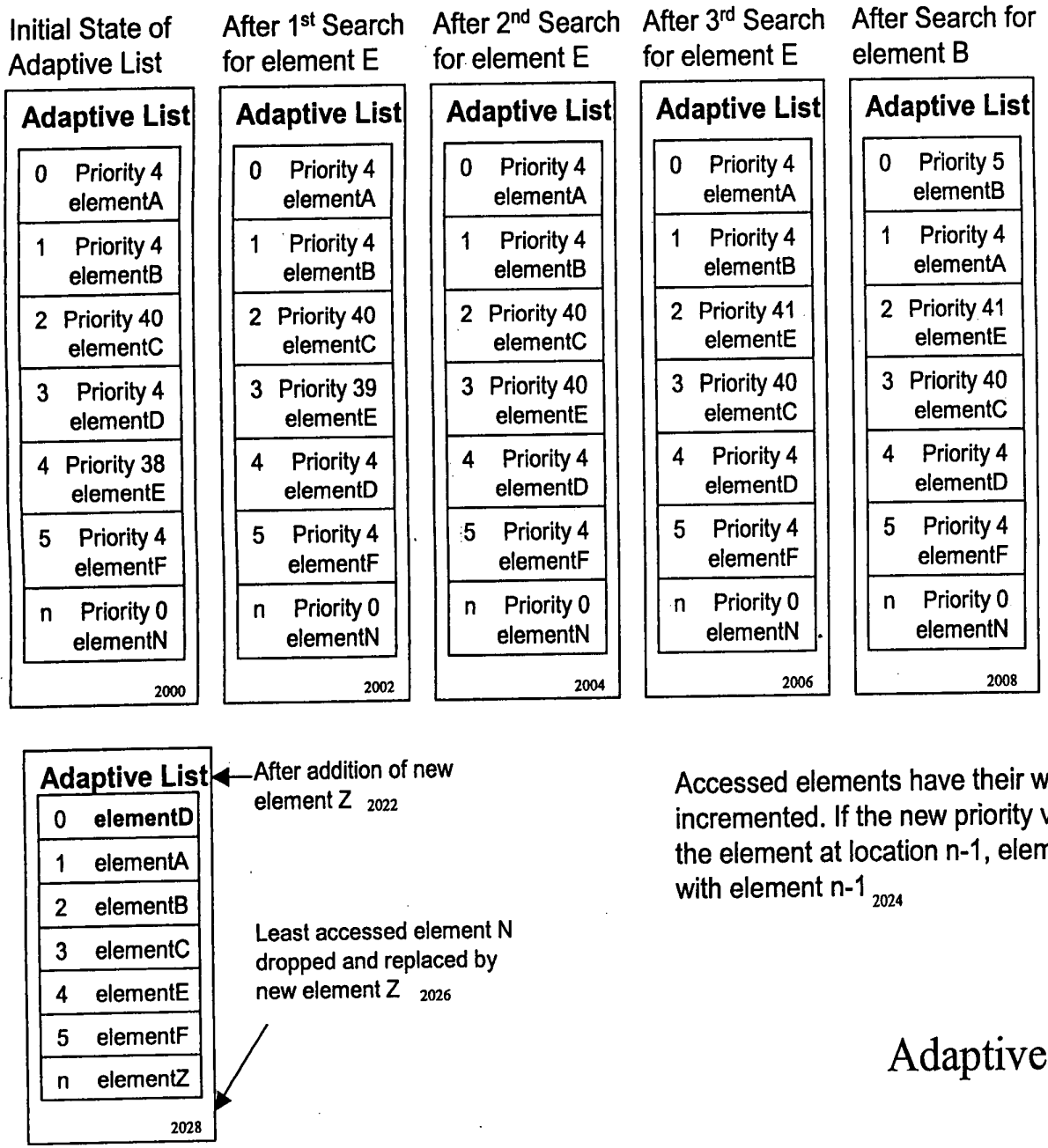
Adaptive Store – Bare Storage
Figure 18

Adaptive Store – Simple Accesses



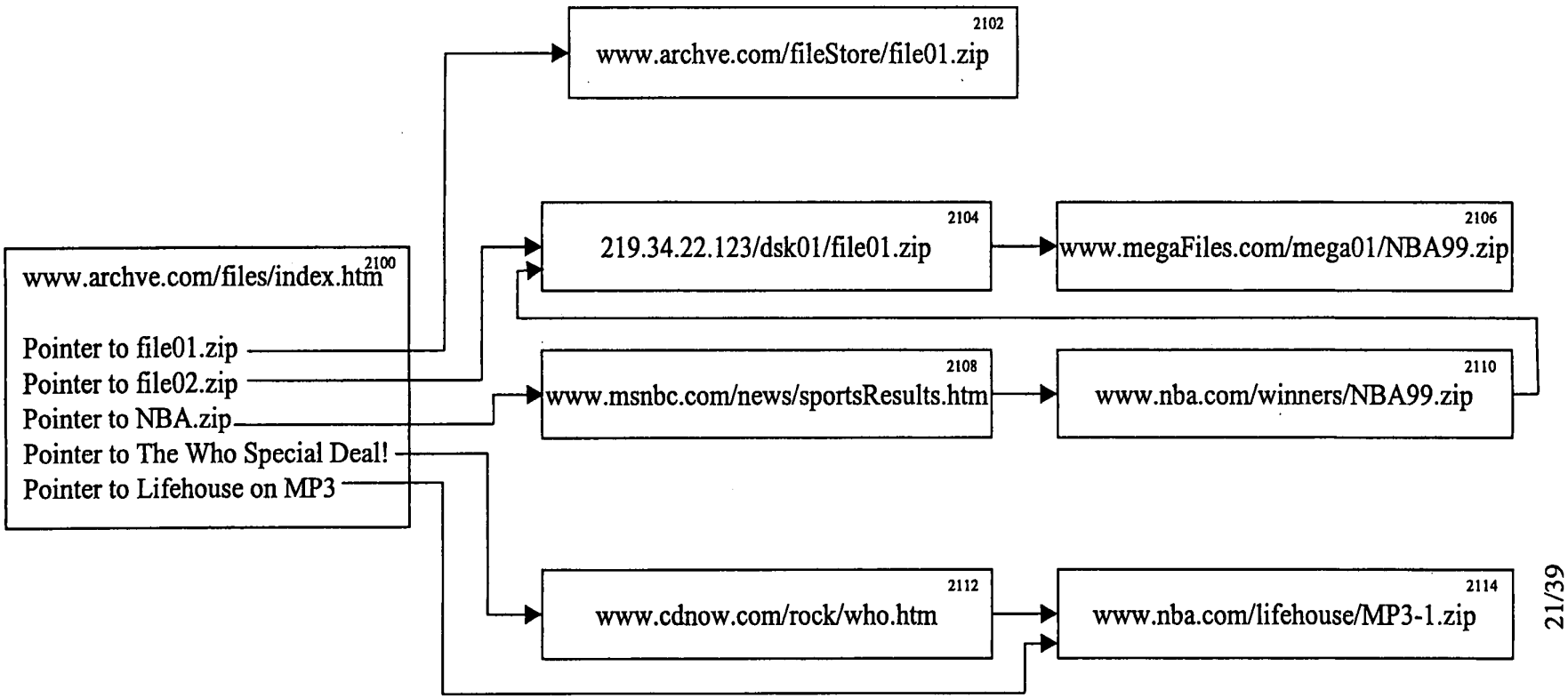
Adaptive Store Simple Access
Figure 19

Adaptive Store – Weighted Accesses



Accessed elements have their weight values incremented. If the new priority value is greater than the element at location n-1, element n is swapped with element n-1 2024

Adaptive Store Weighted Access
Figure 20



Examples of how links to files can span different repositories on a Network such as the WWW and how such links may reference the same file.

URL Chaining and Indirection
Figure 21

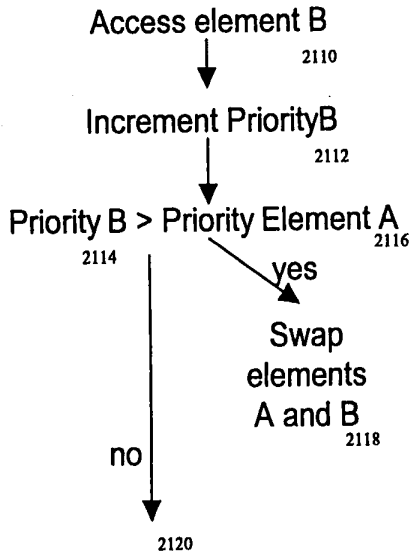
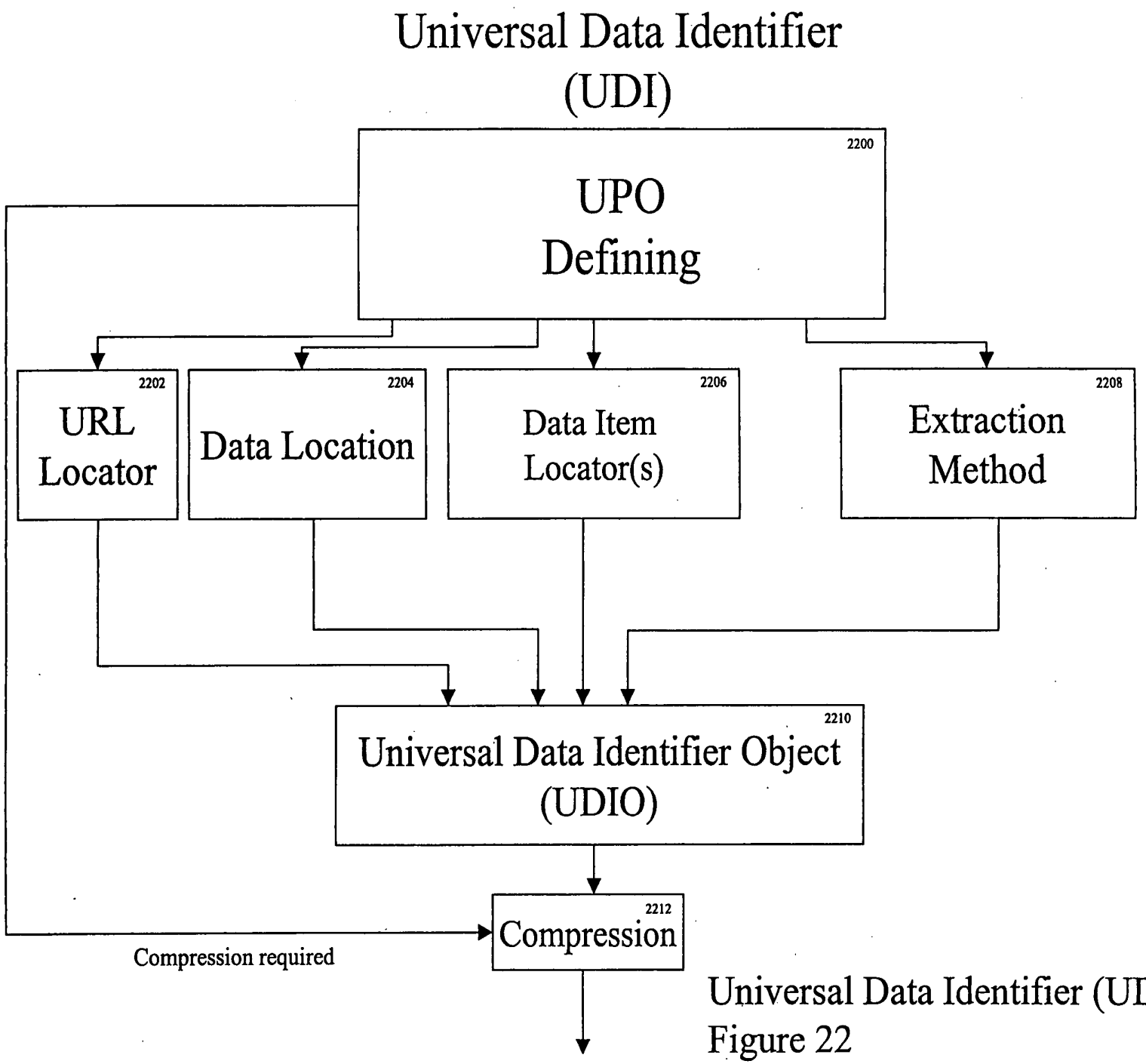
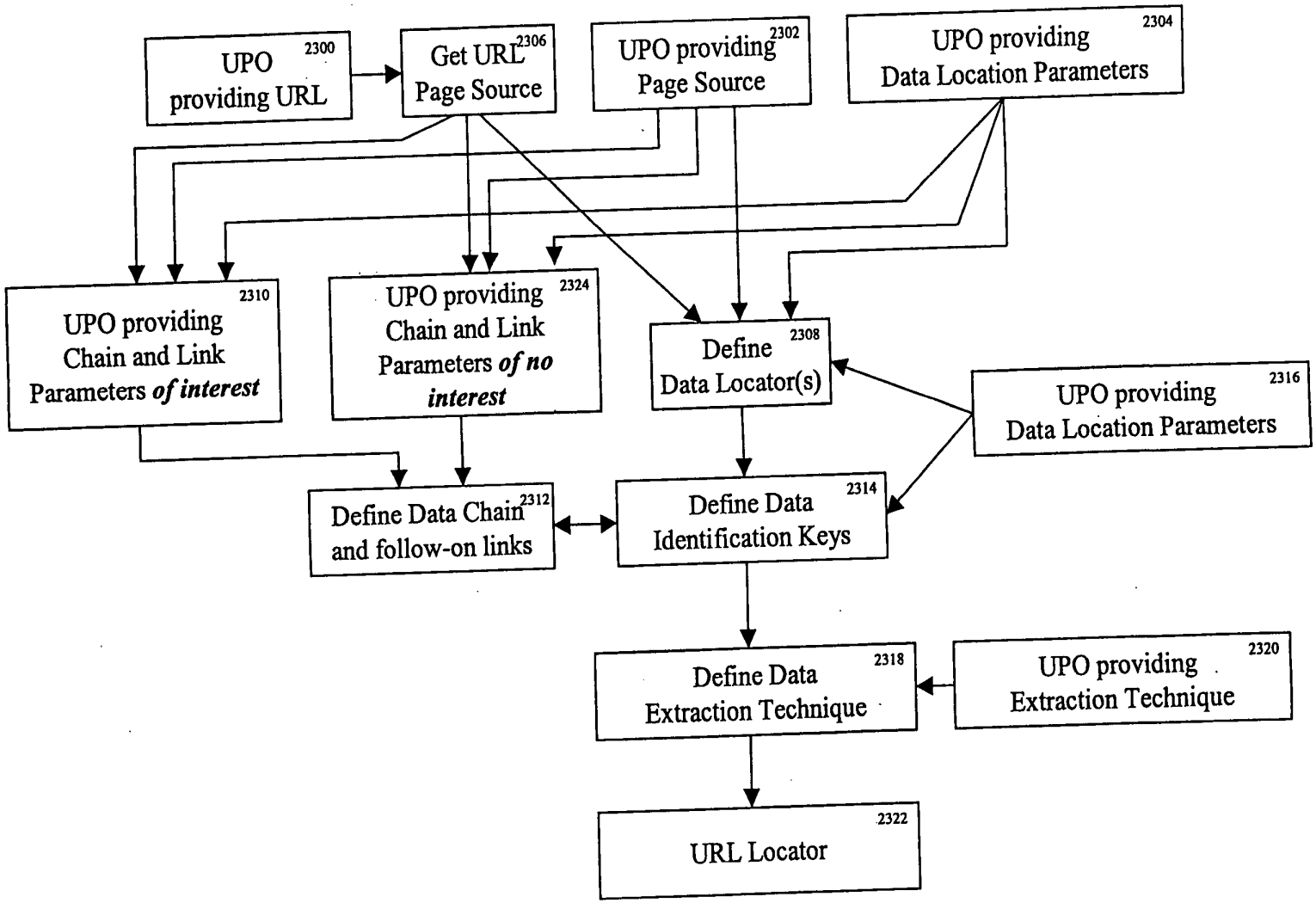

Access element B
2110
Increment PriorityB
2112
Priority B > Priority Element A
2114
2116
yes
Swap
elements
A and B
2118
no
2120

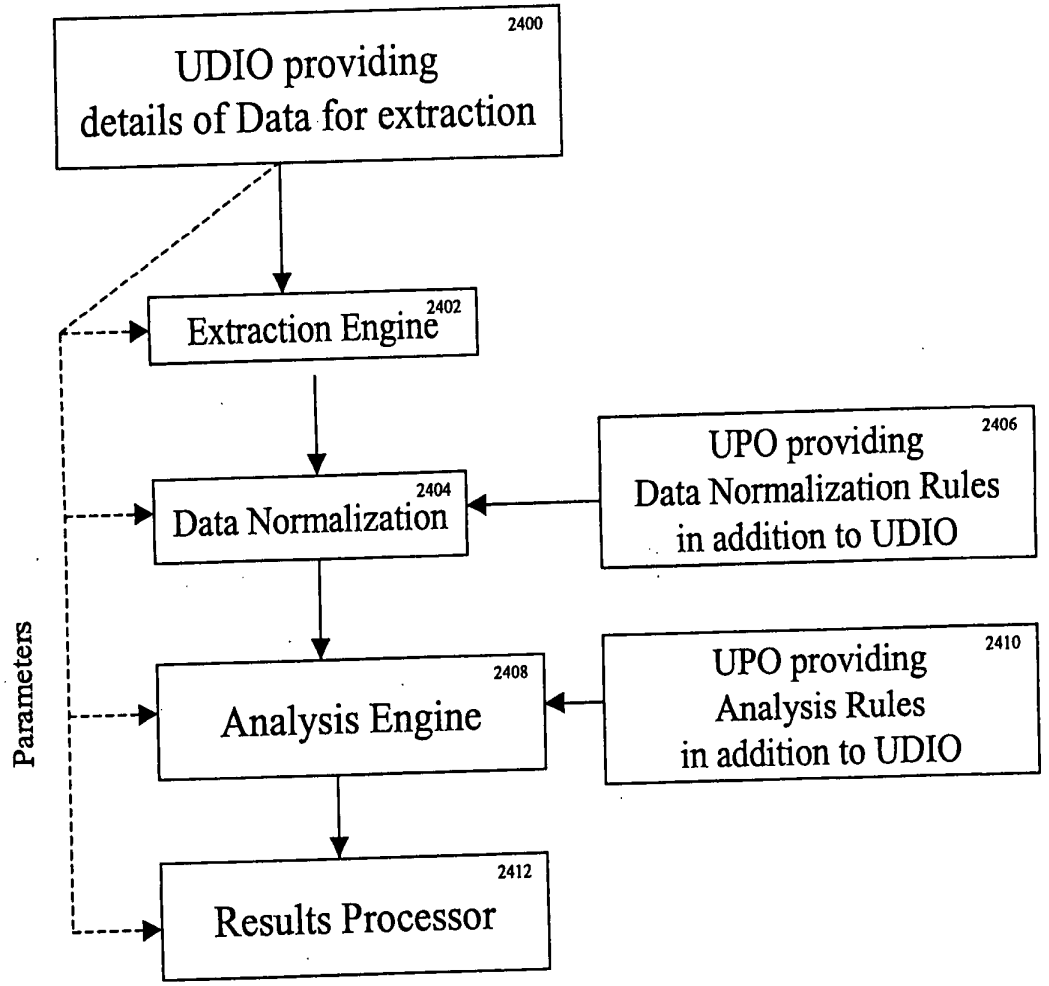
Figure 20-1



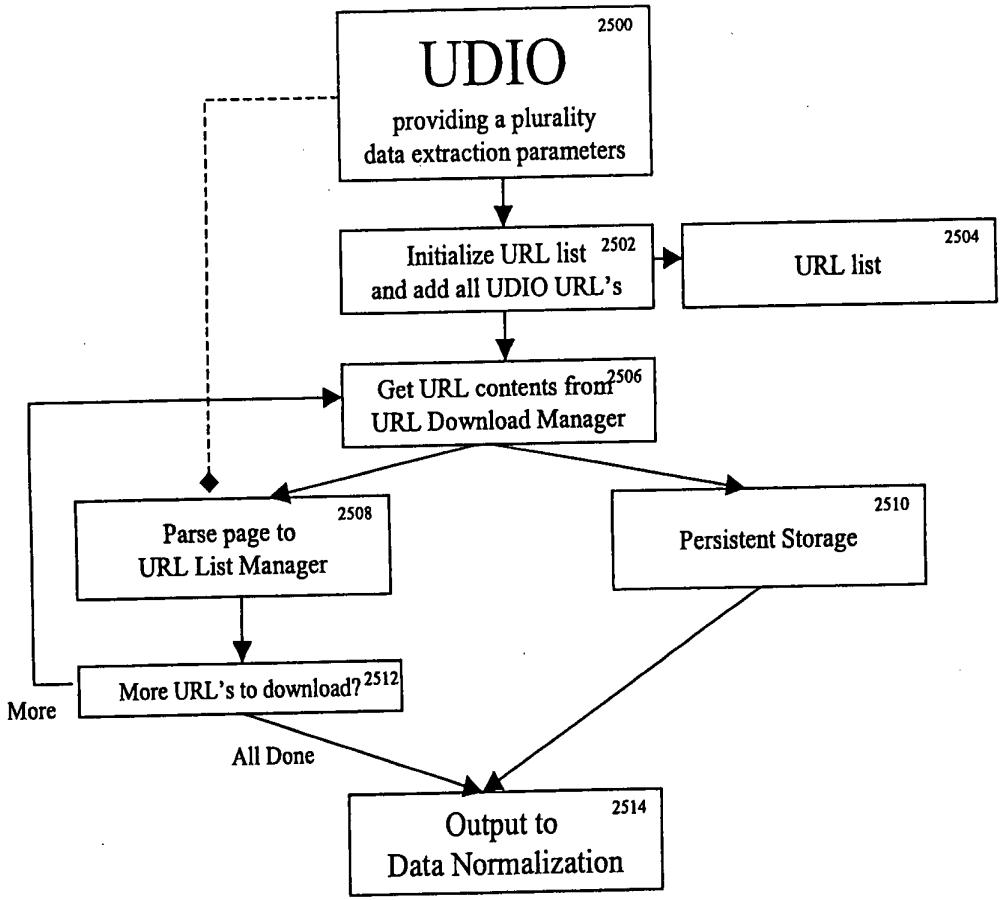
Universal Data Identifier (UDI)
Figure 22



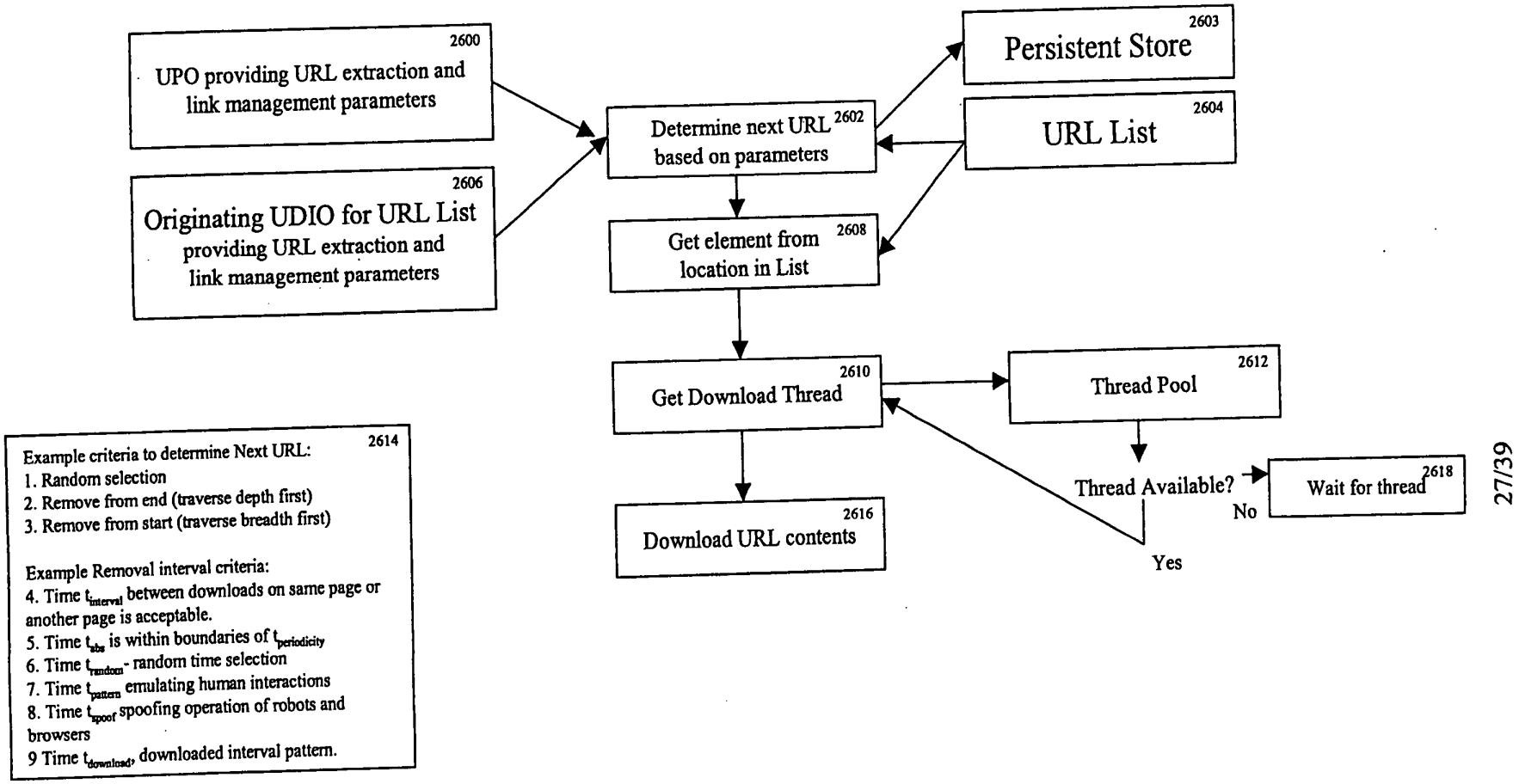
URL Universal Data Identifier
Figure 23



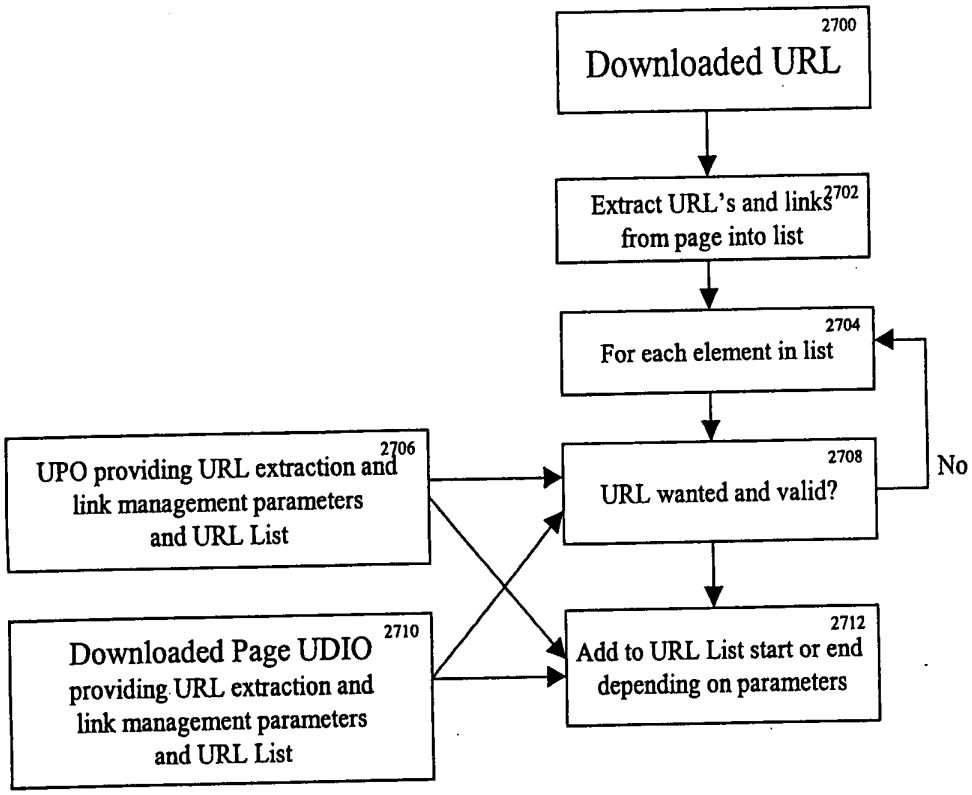
Client Data Extractor (CDE)
Figure 24



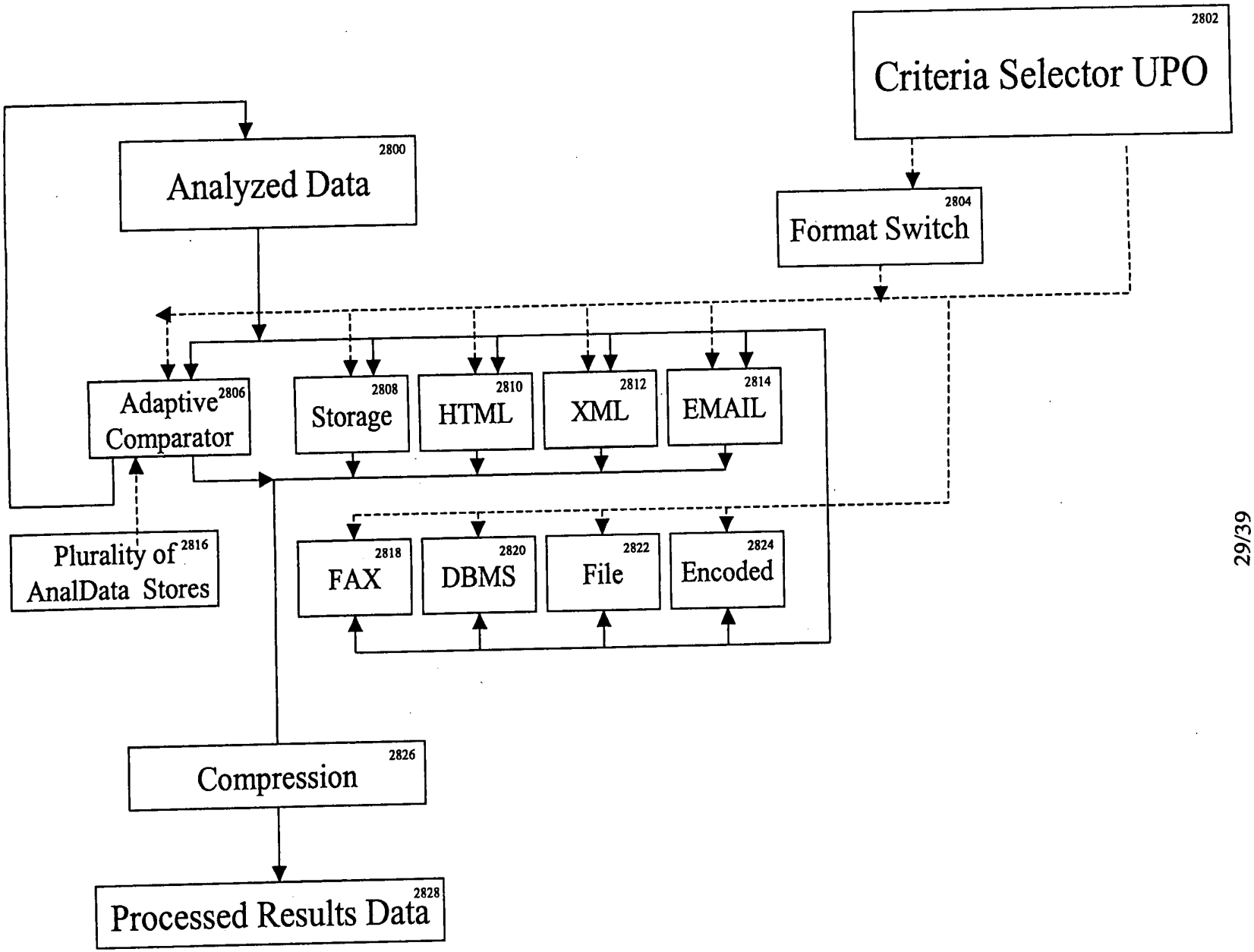
WWW Extraction Engine,
Figure 25



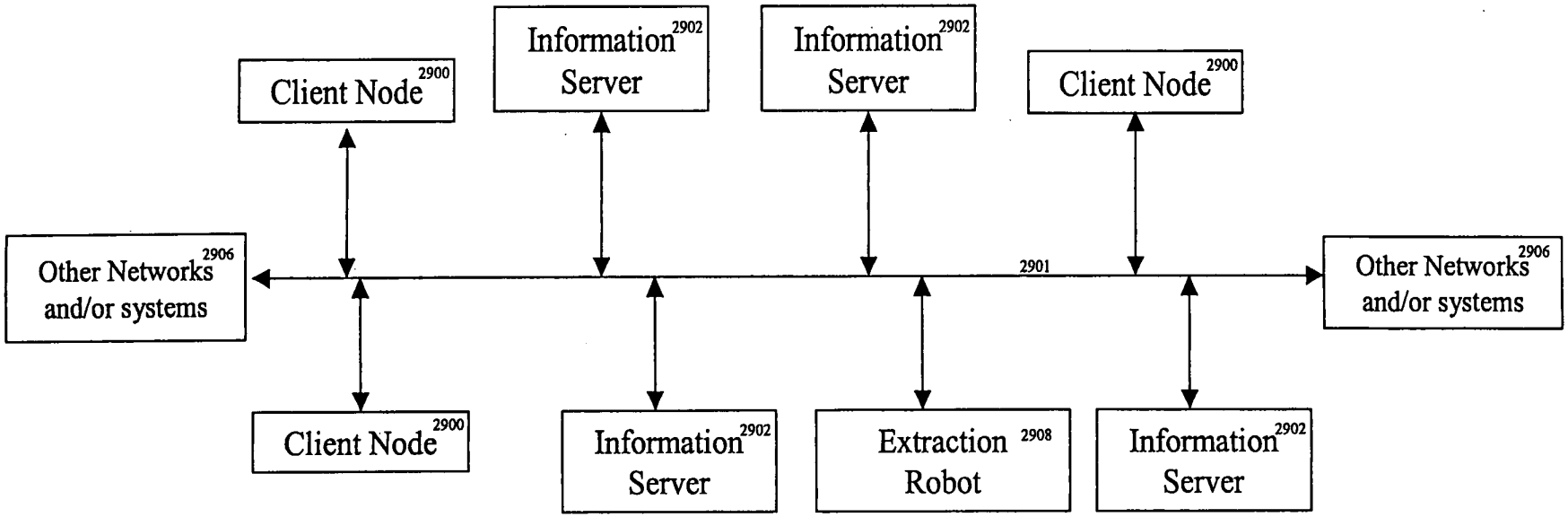
URL Download Manager
Figure 26



URL List Manager
Figure 27



Data Results Processor (DRP)
Figure 28



Examples of information servers being accessed from Human operators and mechanized devices.

Figure 29

WWW Page Load, Processing and Display Times 3002

t^{p}_{text} = Time to obtain, process and display text

t^{p}_{image} = Time to obtain, process and display images

t^{p}_{other} = Time to obtain, process and display all other items

$t^{p}_{total} = t^{p}_{text} + t^{p}_{image} + t^{p}_{other}$ Total time to load, process and display all page items

Human times to access URL from a Displayed Page 3004

$t^{h}_{response}$ = Time to react to and access URL

$t^{h}_{internal}$ = Time for apparatus to respond to URL access

t^{h}_{other} = other miscellaneous times

$t^{h}_{min} = t^{h}_{response} + t^{h}_{other} + t^{h}_{internal}$ Total time to react to and access a URL

t^{h}_{max} = infinite.

Non-Human **times to access URL from a Displayed Page** 3006

t^{n}_{text} = Time to obtain, process and display text

$t^{p}_{internal}$ = Time for apparatus to respond to URL access (very small)

t^{n}_{other} = Time to obtain, process and display all other item (tending to zero)

$t^{n}_{min} = t^{n}_{text} + t^{n}_{internal} + t^{n}_{other}$ Total time to react to and access a URL

t^{n}_{max} = infinite.

Figure 30
Timing Definitions

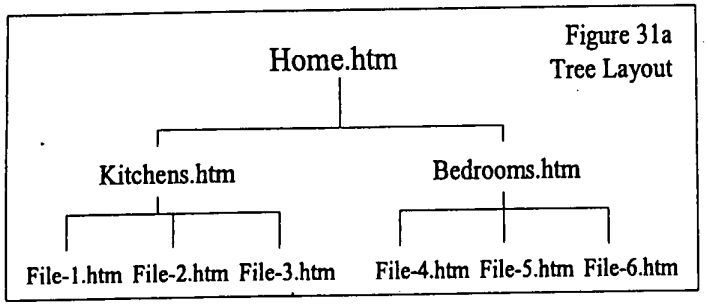


Figure 31a
Tree Layout

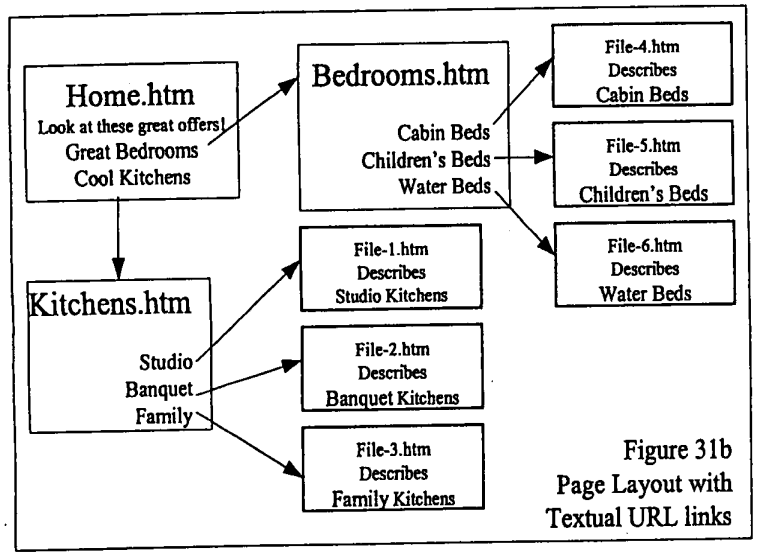


Figure 31b
Page Layout with
Textual URL links

	3110 Requester ID	3112 Data Item ID	3114 Time Stamp	3116 Type of Access
3118	193.133.51.2	/Home.htm	02/14/00 10:43:15223	Read OK
3120	193.133.51.2	/Home/Bedrooms.htm	02/14/00 10:43:17554	Read OK
3122	193.133.51.2	/Home/Bedrooms /File-4.htm	02/14/00 10:43:20332	Read OK
3124	193.133.51.2	/Home/Bedrooms.htm	02/14/00 10:43:21555	Read OK
3126	193.133.51.2	/Home/Bedrooms /File-5.htm	02/14/00 10:43:24676	Read OK
3128	193.133.51.2	/Home/Bedrooms.htm	02/14/00 10:43:26009	Read OK
3130	193.133.51.2	/Home/Bedrooms /File-6.htm	02/14/00 10:43:29876	Read OK
3132	193.133.51.2	/Home.htm	02/14/00 10:43:31000	Read OK
3134	193.133.51.2	/Home/Kitchens.htm	02/14/00 10:43:33442	Read OK
3136	193.133.51.2	/Home/Kitchens /File-1.htm	02/14/00 10:43:35998	Read OK
3138	193.133.51.2	/Home/Kitchens.htm	02/14/00 10:43:37009	Read OK
3140	193.133.51.2	/Home/Kitchens /File-2.htm	02/14/00 10:43:40030	Read OK
3142	193.133.51.2	/Home/Kitchens.htm	02/14/00 10:43:41993	Read OK
3144	193.133.51.2	/Home/Kitchens /File-3.htm	02/14/00 10:43:43554	Read OK

Example of server access by a *human* operator

Figure 31c

	3150 Requester ID	3152 Data Item ID	3154 Time Stamp	3156 Type of Access
3158	193.133.51.2	/Home.htm	02/14/00 10:43:15000	Read OK
3160	193.133.51.2	/Home/Bedrooms.htm	02/14/00 10:43:15550	Read OK
3162	193.133.51.2	/Home/Kitchens.htm	02/14/00 10:43:15553	Read OK
3164	193.133.51.2	/Home/Bedrooms/File-4.htm	02/14/00 10:43:16000	Read OK
3166	193.133.51.2	/Home/Bedrooms/File-5.htm	02/14/00 10:43: 16005	Read OK
3168	193.133.51.2	/Home/Bedrooms/File-6.htm	02/14/00 10:43: 16010	Read OK
3170	193.133.51.2	/Home/Kitchens/File-1.htm	02/14/00 10:43: 16020	Read OK
3172	193.133.51.2	/Home/Kitchens/File-2.htm	02/14/00 10:43: 16030	Read OK
3174	193.133.51.2	/Home/Kitchens/File-3.htm	02/14/00 10:43: 16040	Read OK

Example of server access by a *non-human* operator such as an Extraction Robot

Figure 31d

Figure 31
Page hierarchy with textual URL's

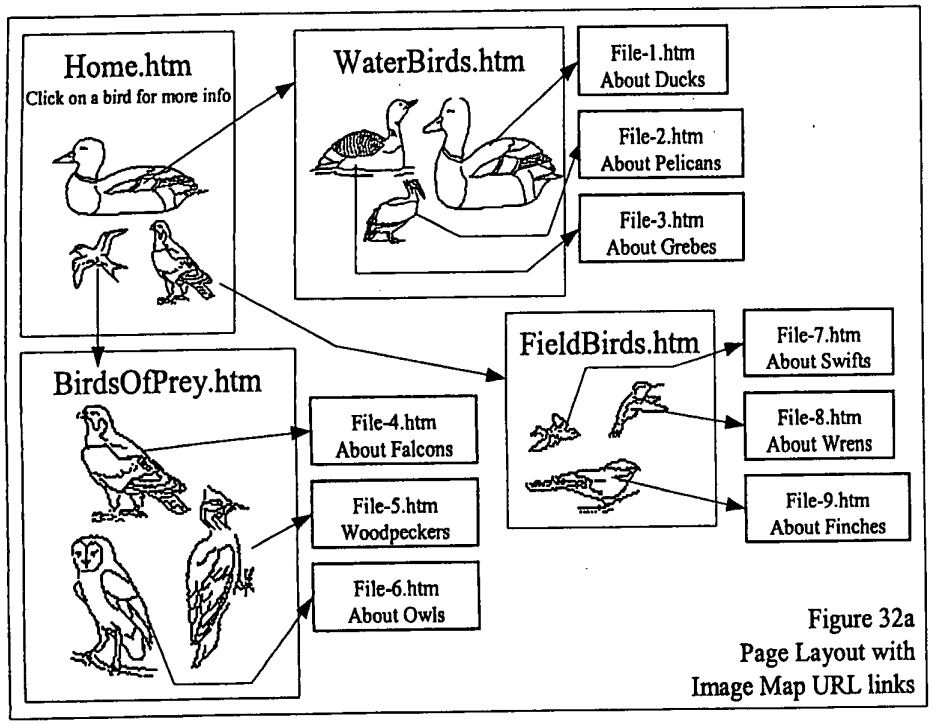


Figure 32a
Page Layout with
Image Map URL links

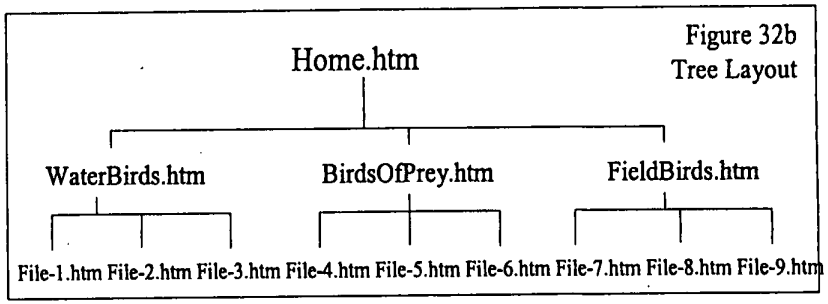


Figure 32b
Tree Layout

Figure 32
Page Hierarchy with Image Maps

Reference times to access a URL

3300 $t^{ref}_{response}$ = Time to react to and access URL

3302 $t^{ref}_{internal}$ = Time for apparatus to respond to URL access

3304 t^{ref}_{other} = other miscellaneous times

3306 $t^{ref}_{min} = t^{h}_{response} + t^{h}_{internal} + t^{h}_{other}$ Total time to react to and access a URL

3308 t^{ref}_{max} = infinite.

Figure 33a

Timing Definitions

3310 $\Delta hit_n = hit_{n+1} - hit_n$ Time difference between 2 hits

3312 $t_{hit\_av} = \sum_{n0}^{n} \Delta hit_n / n$ Average time for hits n_0 to n

3314 $t^{no \rightarrow n}_{hit\_min}$ Minimum hit time for hits n_0 to n

3316 $t^{no \rightarrow n}_{hit\_max}$ Maximum hit time for hits n_0 to n

hits n_0 is the first hit in a sequence, eg: B320
hit n is the last hit in a sequence, eg: B344

Figure 33b

Human Signature Definitions

3330 $t^{h\_sig}_{min} = t^{hit}_{} - t^{ref}_{min}$ Difference between hit time and reference minimum value

3332 $t^{h\_sig}_{max} = t^{hit}_{} - t^{ref}_{max}$ Difference between hit time and reference maximum value

3334 $t^{h\_sig}_{av} = t^{hit}_{} - t^{ref}_{av}$ Difference between hit time and reference average value

3336 $t^{h\_sig}_{delta} = (t^{h\_sig}_{min} + t^{h\_sig}_{av} + t^{h\_sig}_{max}) / 3$ Average for all human signature values

Figure 33c

Figure 33
Signature Calculations

Signature Proximity Terms

3400 $t_{min}^{prox} = t_{min}^{h\_sig} - t_{min}^{r\_ref}$ Difference between human minimum signature and robot reference minimum values

3402 $t_{av}^{prox} = t_{av}^{h\_sig} - t_{av}^{r\_ref}$ Difference between human average signature and robot reference average values

3404 $t_{max}^{prox} = t_{max}^{h\_sig} - t_{max}^{r\_ref}$ Difference between human maximum signature and robot reference maximum values

3406 $t_{min}^{prob} = t_{min}^{prox} \rightarrow t_{min}^{r\_ref}$ higher probability of robot hit. Decreasing positive values and increasing negative values indicate higher probability.

3408 $t_{av}^{prob} = t_{av}^{prox} \rightarrow t_{av}^{r\_ref}$ higher probability of robot hit. Decreasing positive values and increasing negative values indicate higher probability.

3410 $t_{max}^{prob} = t_{av}^{prox} \rightarrow t_{max}^{r\_ref}$ higher probability of robot hit. Decreasing positive values and increasing negative values indicate higher probability.

t^{r_ref} denotes a robot (ie non-human) *hit reference* term

t^{h_ref} denotes a human *hit reference* term

Figure 34a

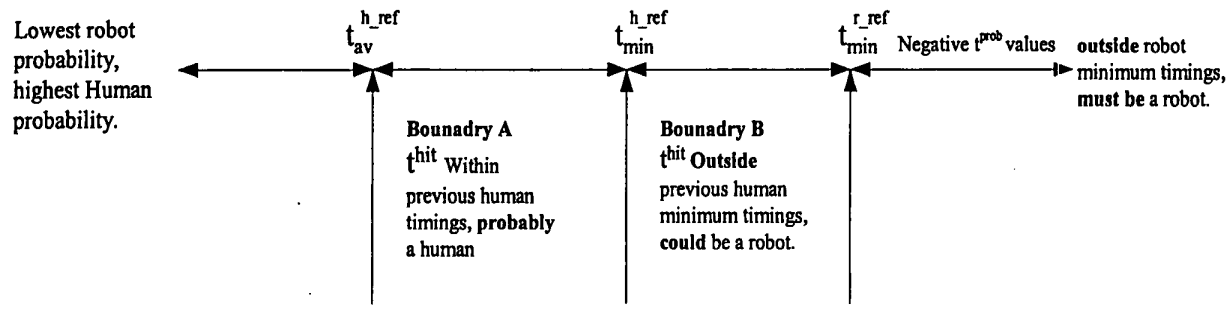


Figure 34b

Figure 34
Signature Proximity Terms

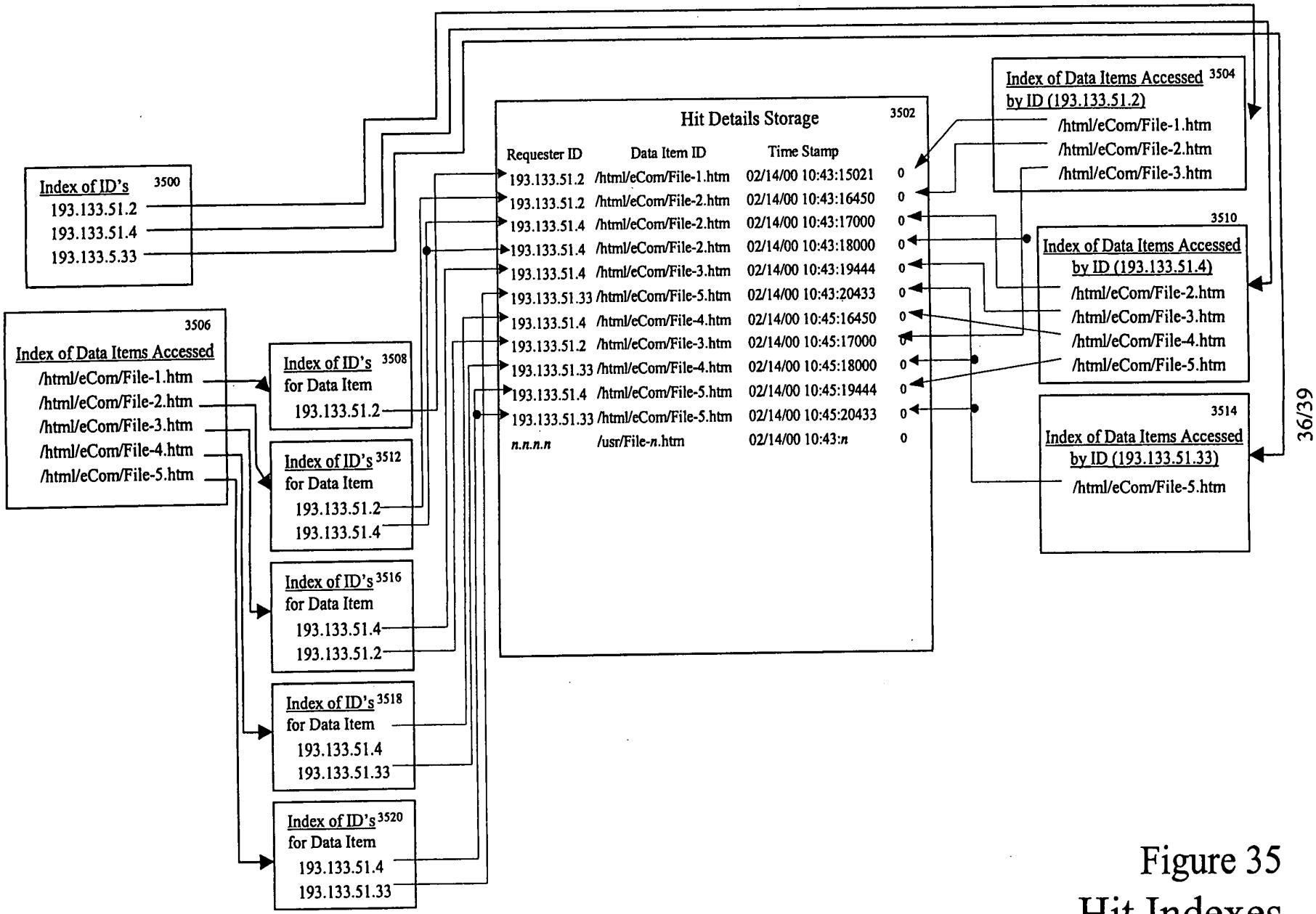
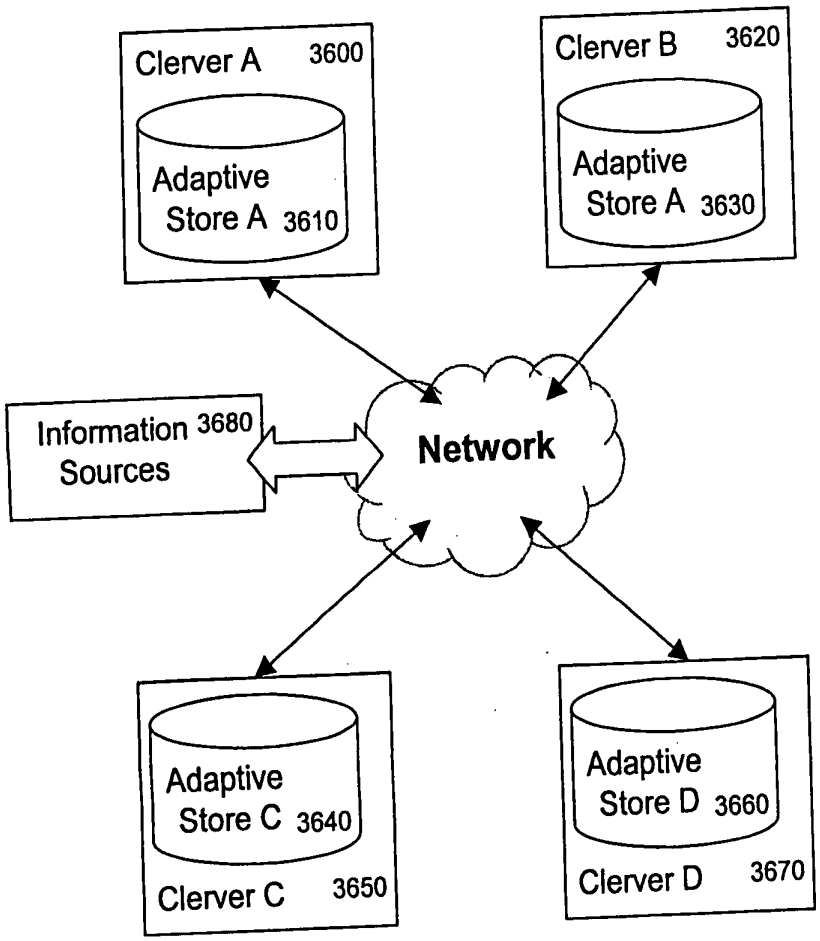
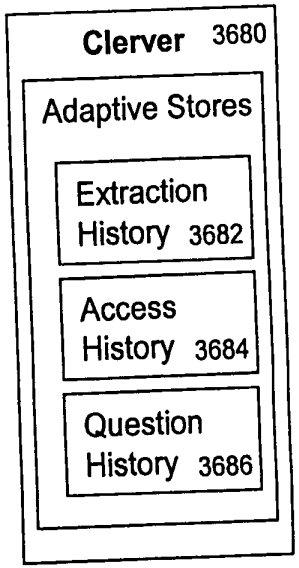
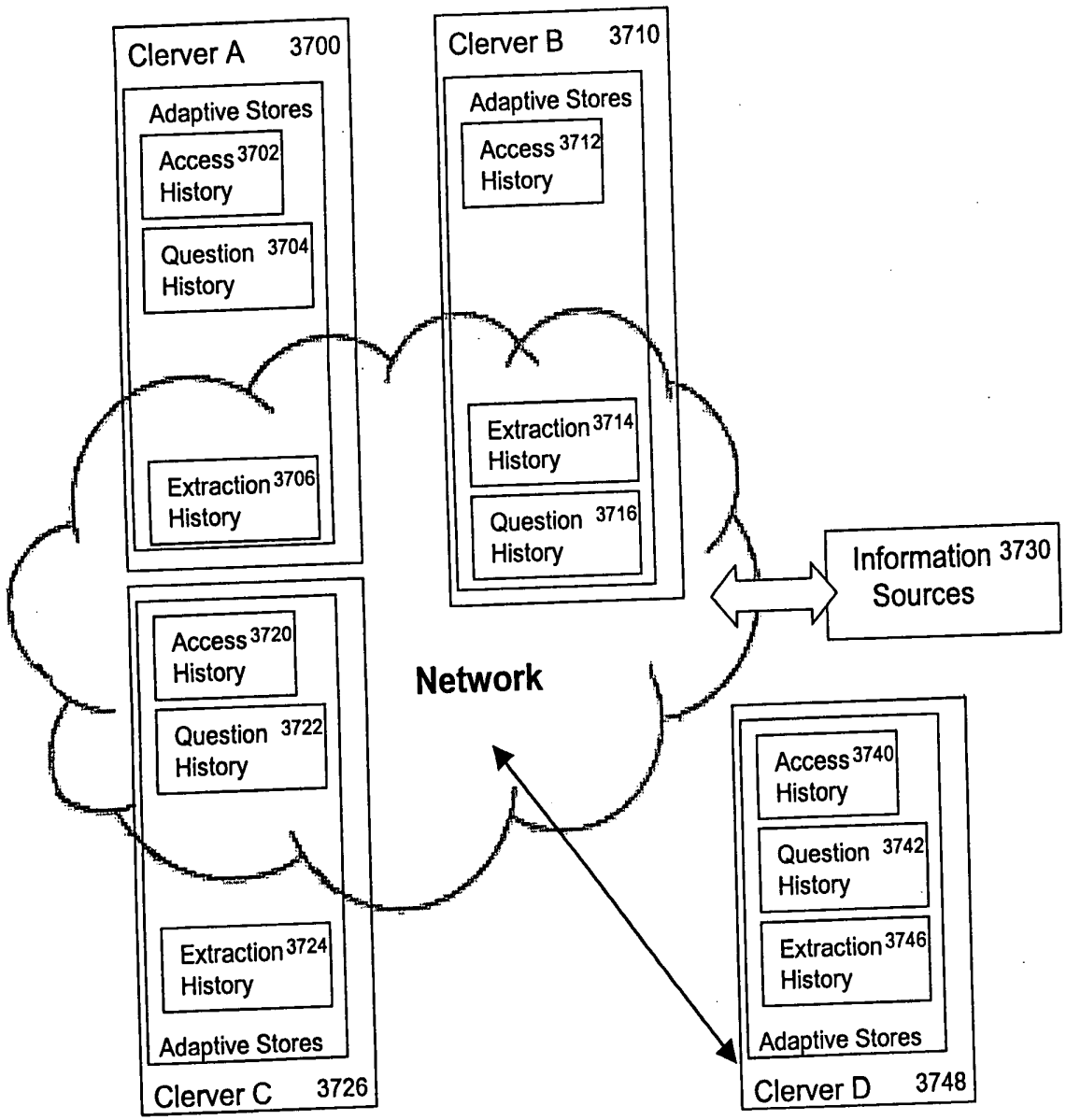


Figure 35
Hit Indexes

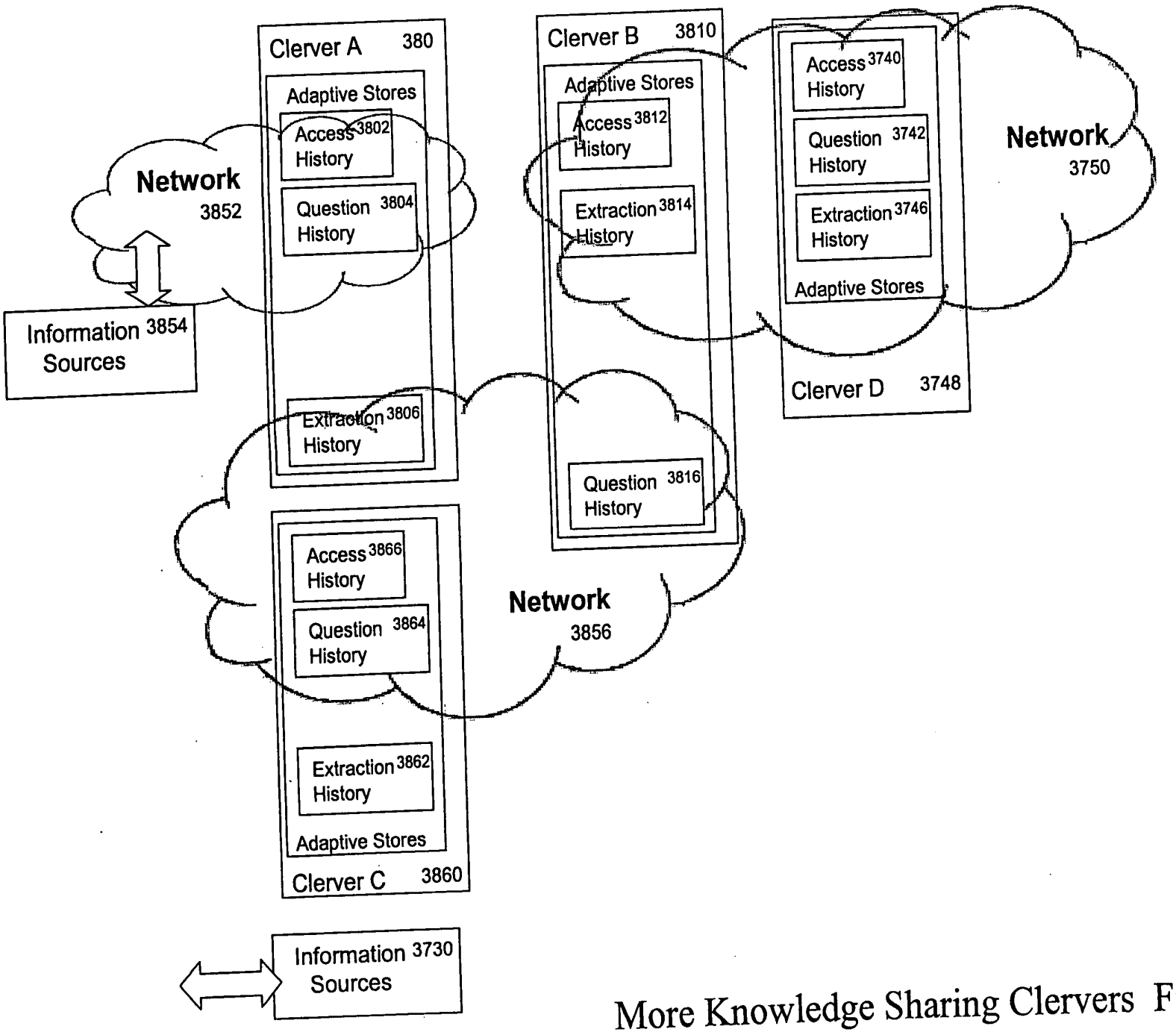




Basic Clerver
Figure 36



Knowledge Sharing Clervers Figure 37



More Knowledge Sharing Clervers Figure 38